

RFUシグナル強度と 正規・対数正規・ポアソン対数正規・二項 分布

～ノイズとシグナル、ホモとヘテロ～

法数学勉強会

6月特別回：法医学会@高知

2015/06/10

京都大学(医) 統計遺伝学分野

山田 亮

今日の主な内容

- 真のシグナルとノイズとの識別
- ヘテロ接合体の2アレルのシグナルのバラツキ

法数学勉強会 例会@京都大学

2015/04/25

との重複があります..

勉強会会場
京大基礎医学
記念講堂



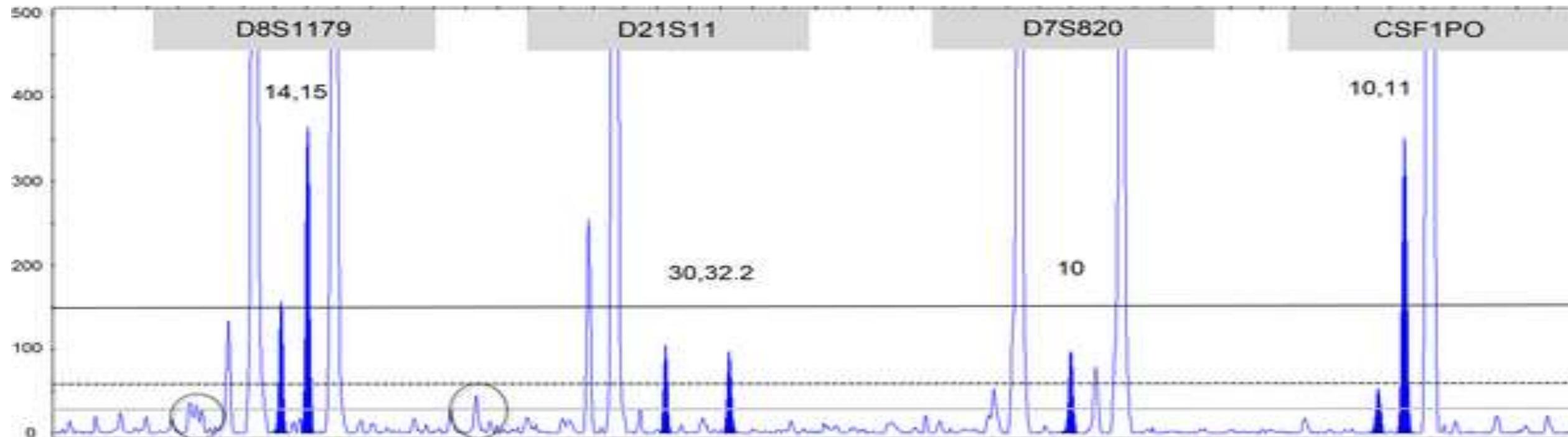
["http://www.genome.med.kyoto-u.ac.jp/wiki_tokyo/index.php講義・資料など：年度別"](http://www.genome.med.kyoto-u.ac.jp/wiki_tokyo/index.php講義・資料など：年度別)

から過去の勉強会の資料(2015/04/25分を含む)が見られます

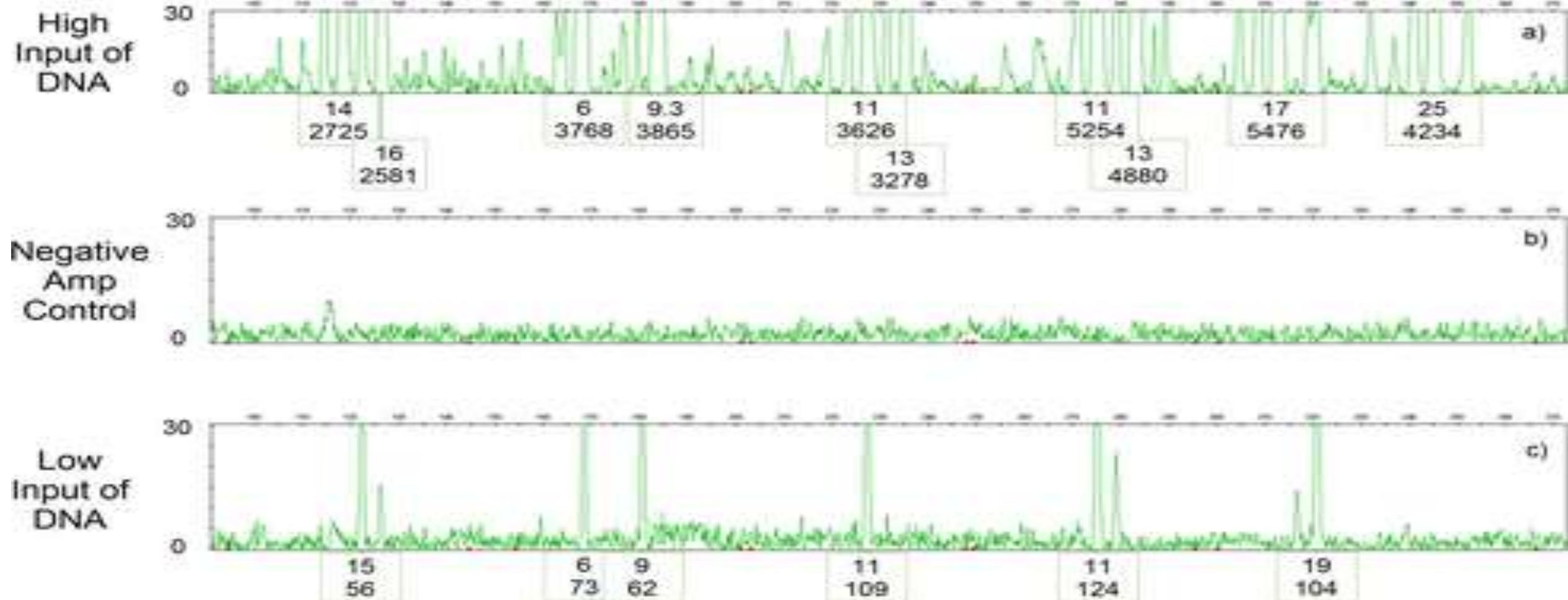


ノイズ

RFU (Relative Fluorescence Units)とノイズ



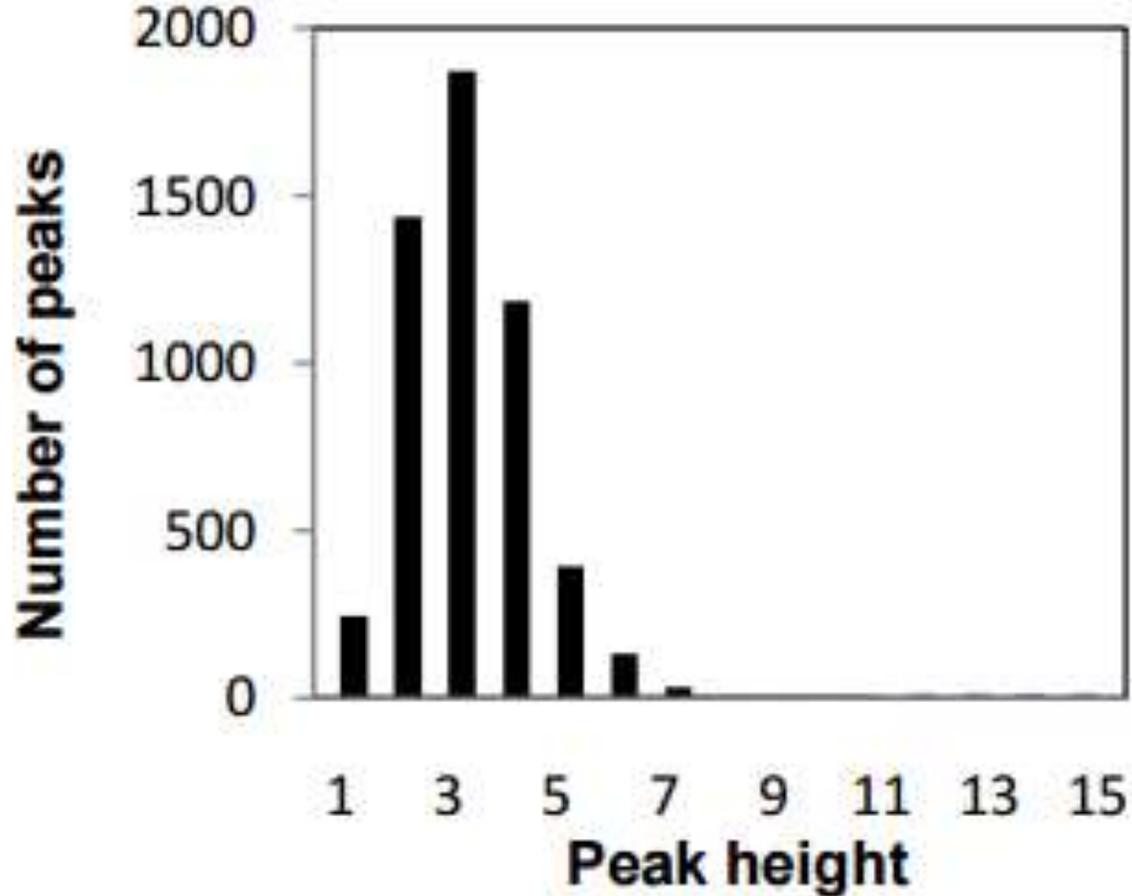
多量・少量DNAの場合のノイズと、ネガコンのノイズ



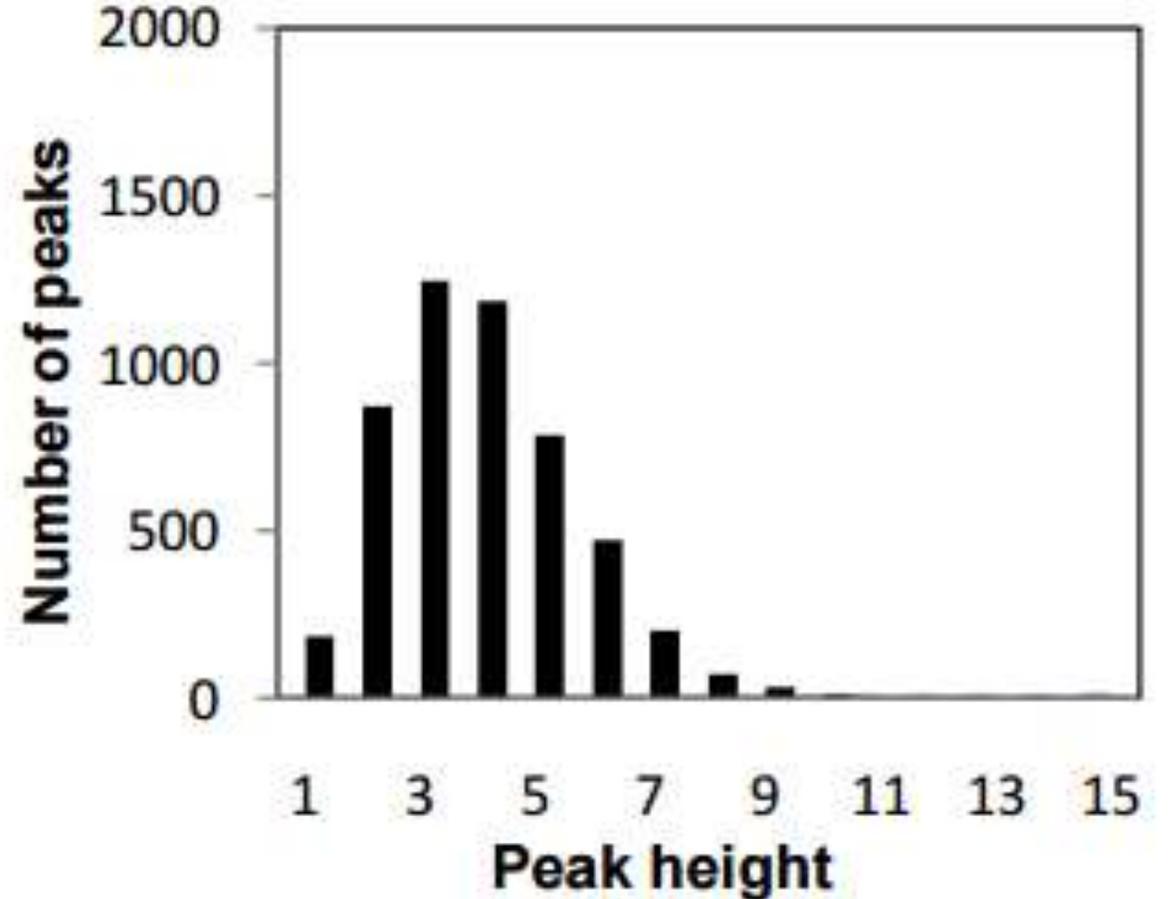
ノイズピーク高の分布

0は「ピークではない」からカウントされない

DNAなしでPCR実施

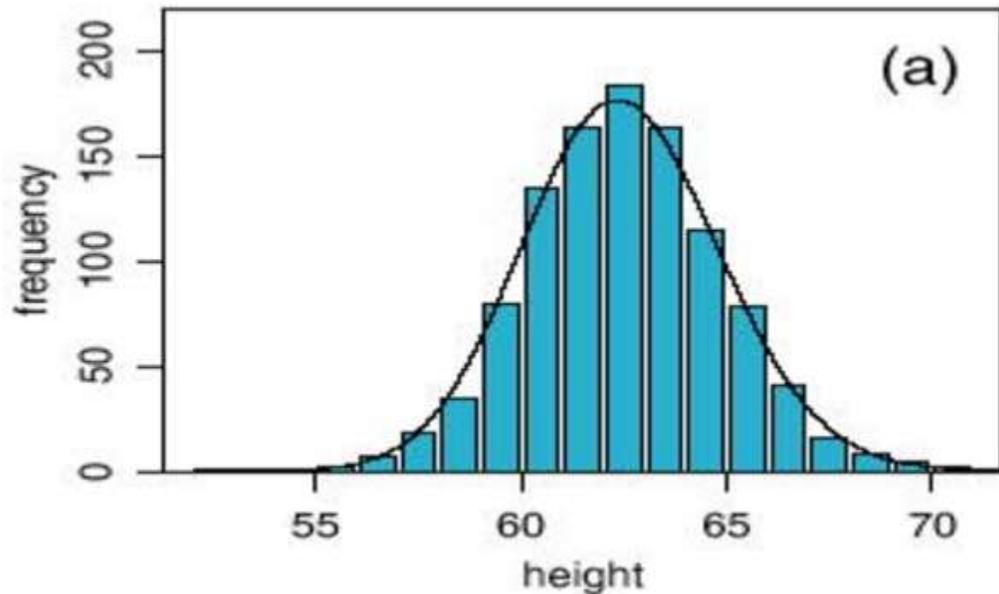


DNAあり、PCRなし



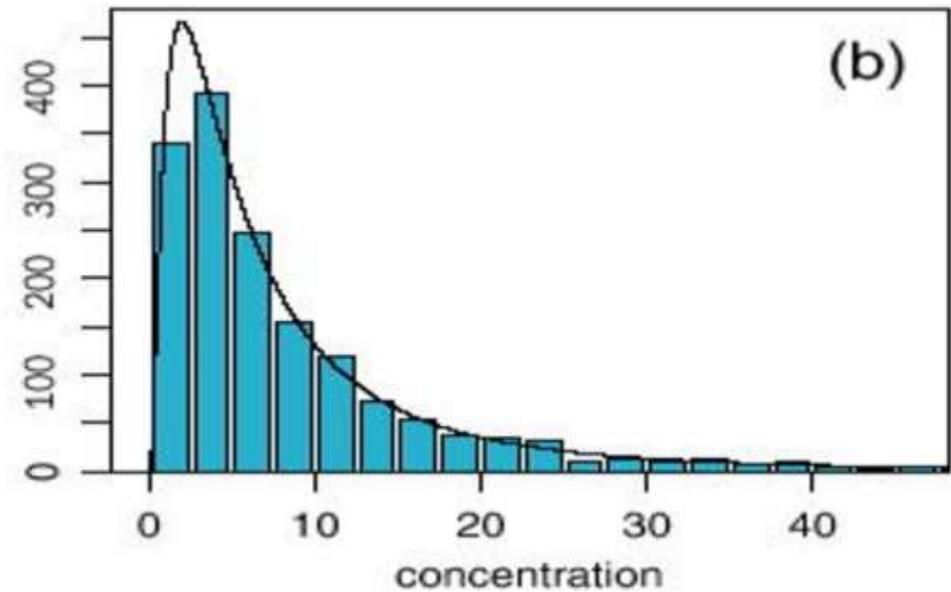
正規分布

- 左右対称
- 正・負

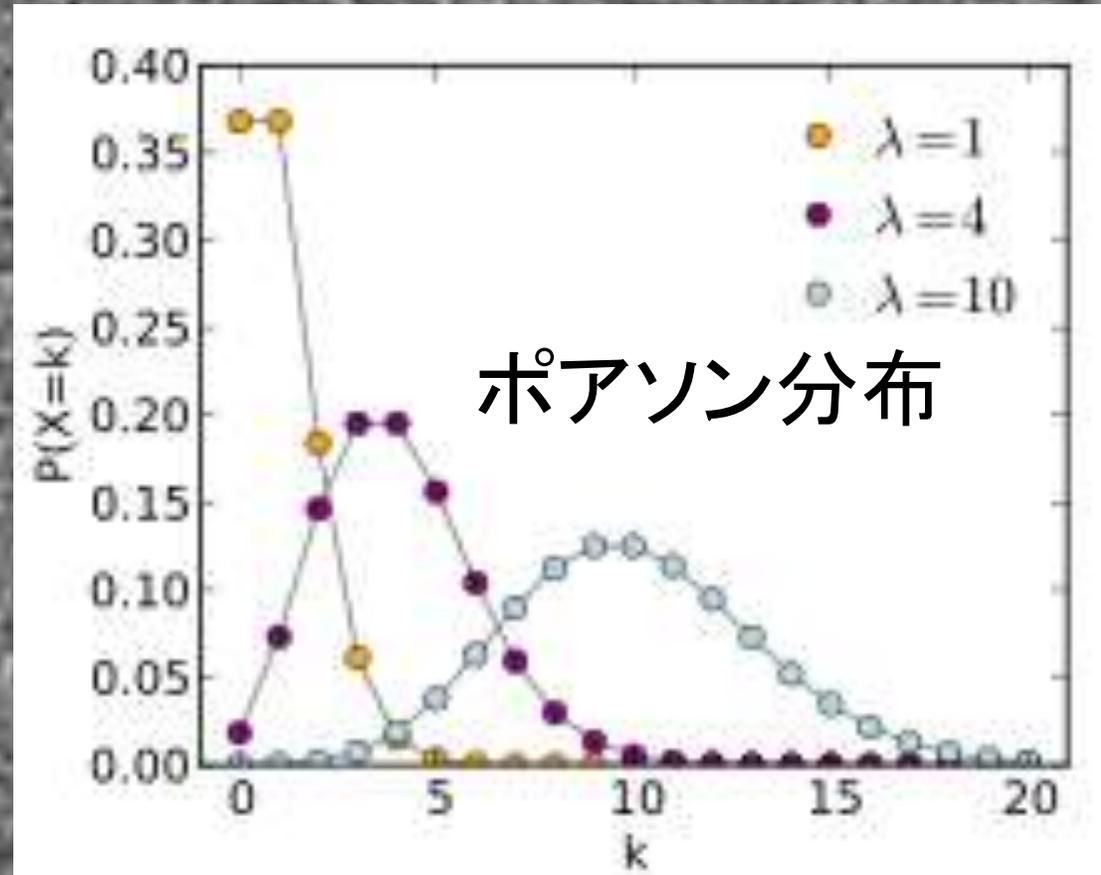
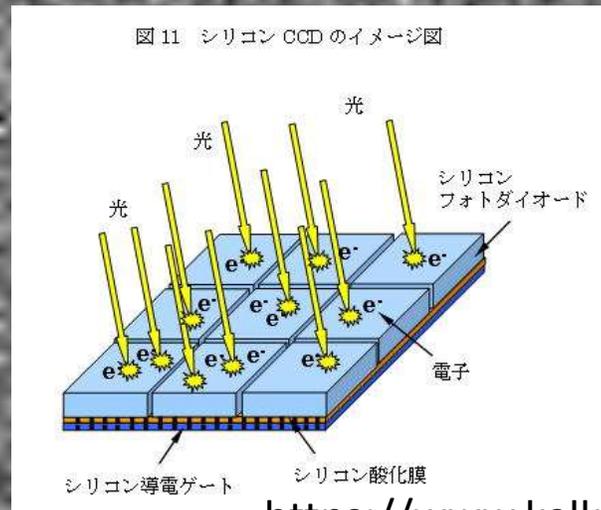


対数正規分布

- 左右非対称
- 正のみ



CCDカメラのノイズ

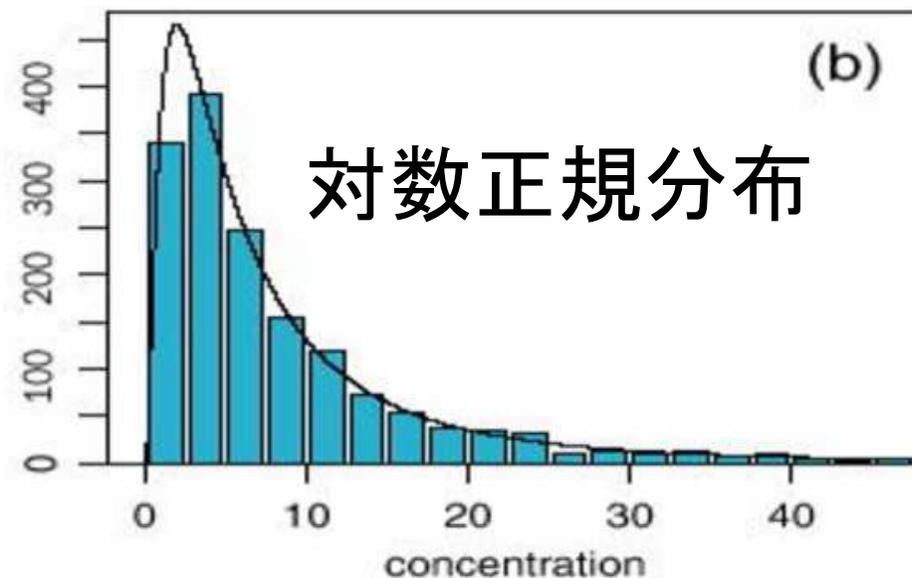
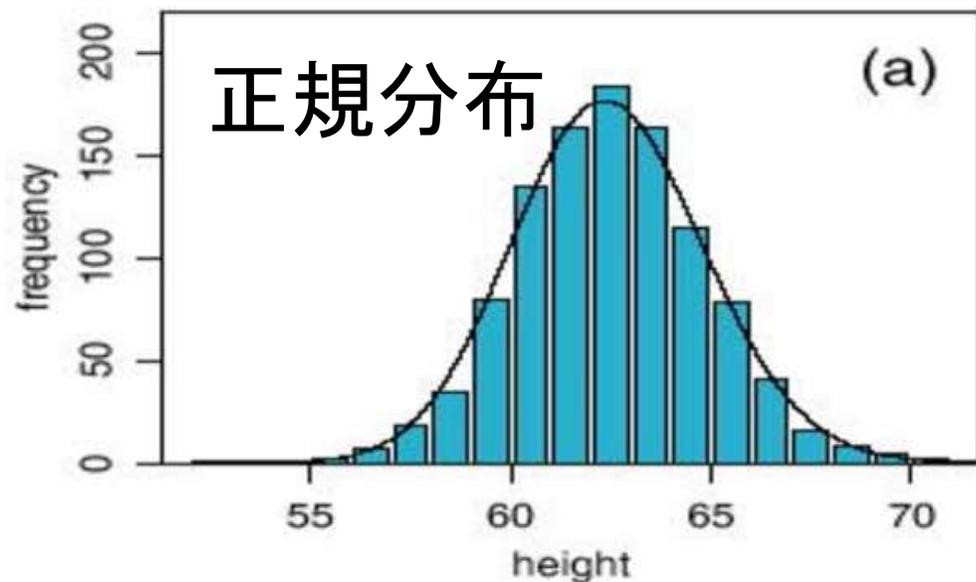


http://qsimaging.com/ccd_noise.html

<https://www.kelk.co.jp/useful/netsuden9.html>

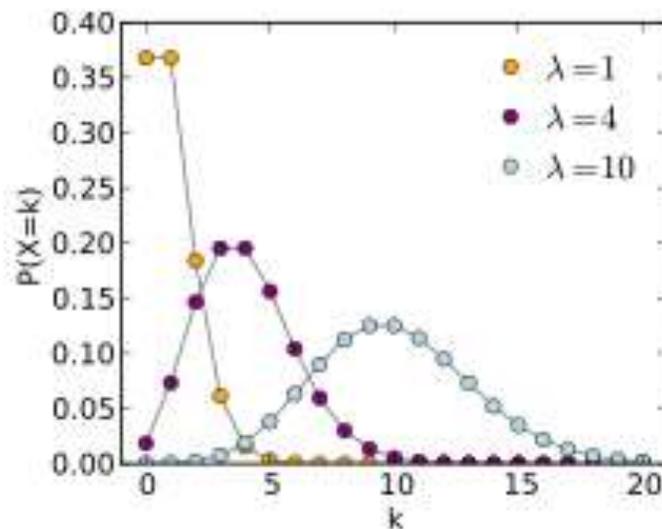
CCDカメラのノイズ ショット・ノイズ





簡単なまとめ
ノイズの分布

非対称
離散



ポアソン分布

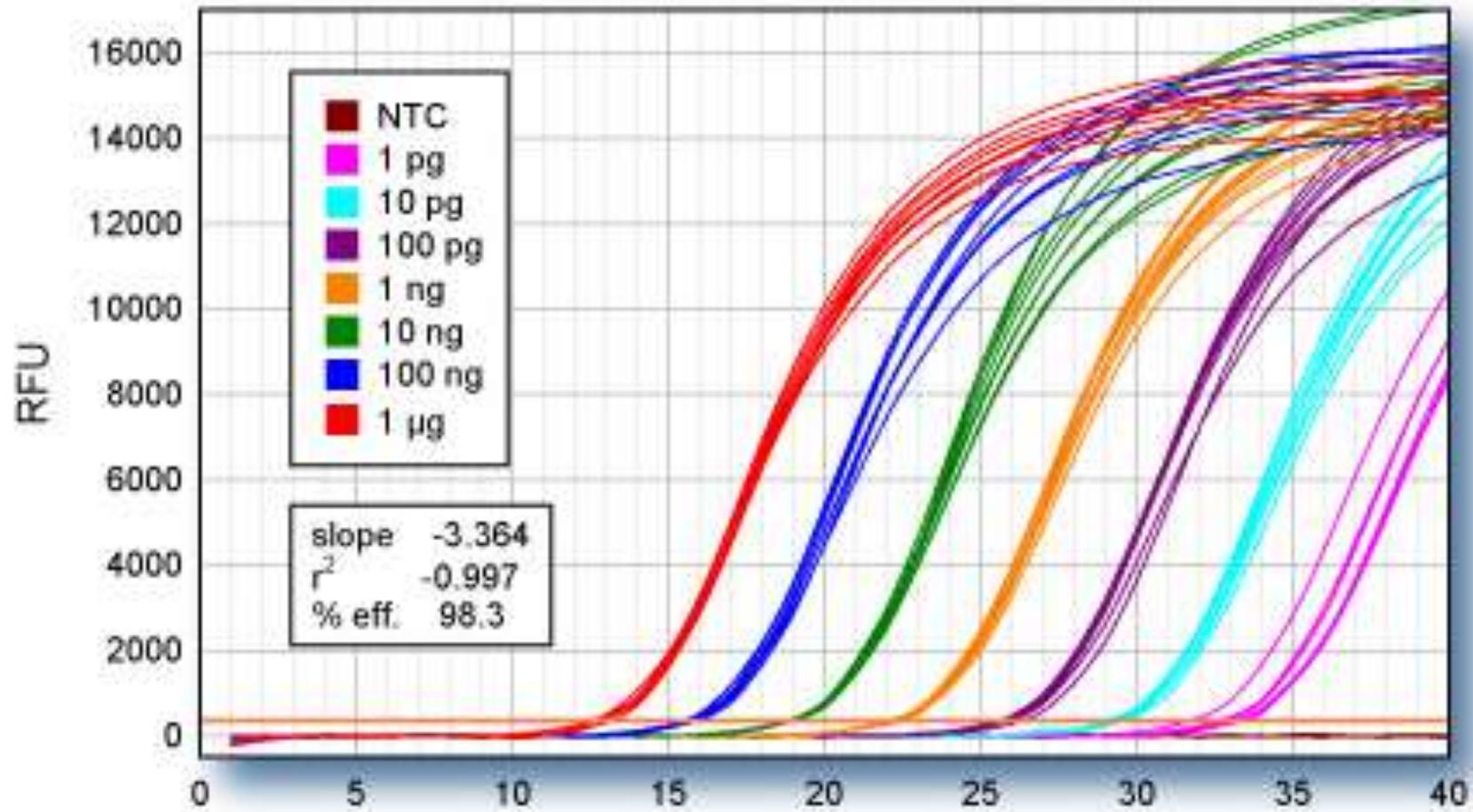
真のシングナナル

ホモ接合体のピーク高はヘテロ接合体の
ピーク高の、だいたい、半分

ホモ接合体のピーク高はヘテロ接合体のピーク高の、だいたい、半分

- ヘテロ接合体のピークの方がノイズに埋もれやすい
- ヘテロ接合体なのに、片方のピークだけが見えることがある
- ピーク高の定量性の話
 - エリア面積のことは、今回は忘れることにします

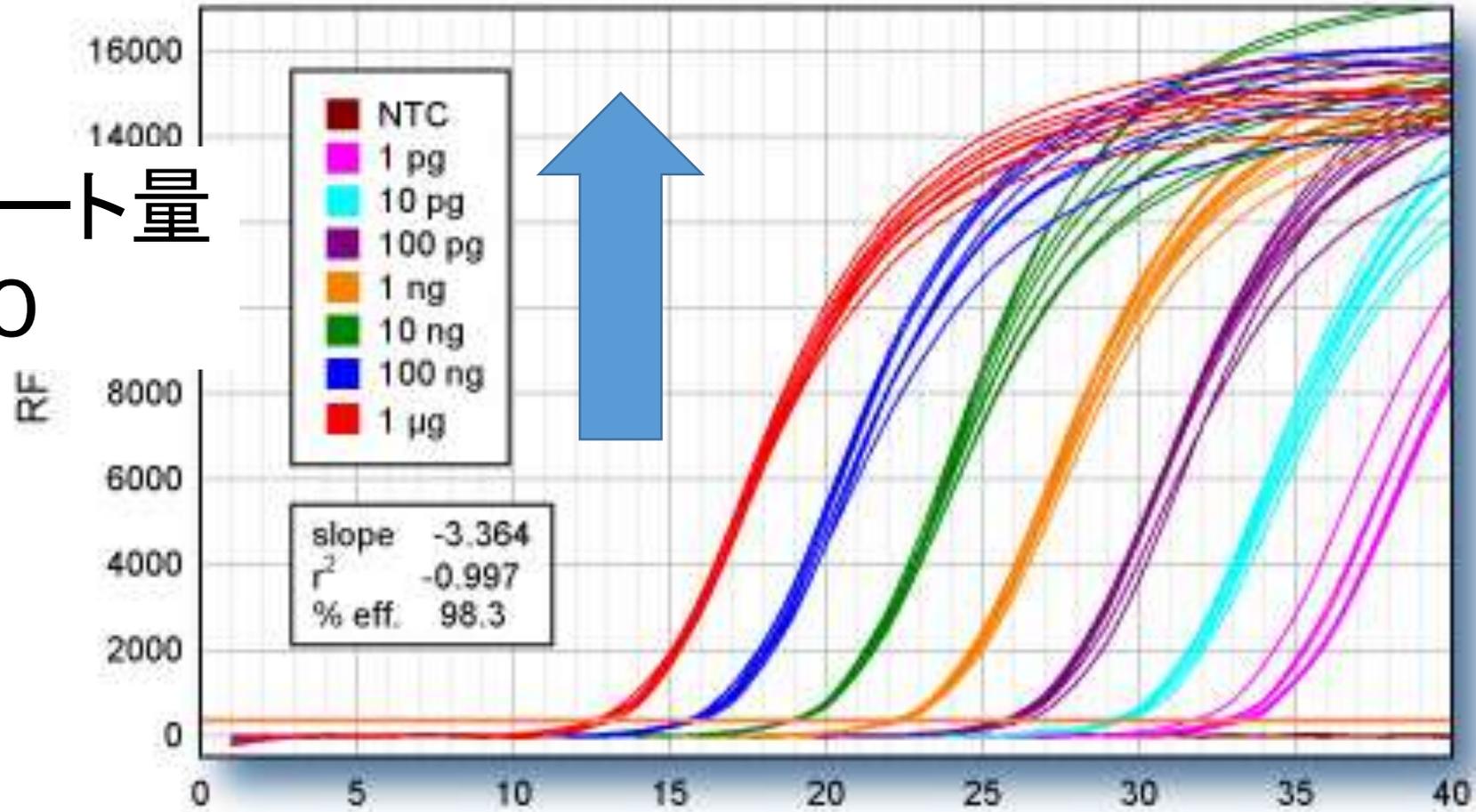
定量的PCR 発現量解析(mRNA)の場合



サイクル数

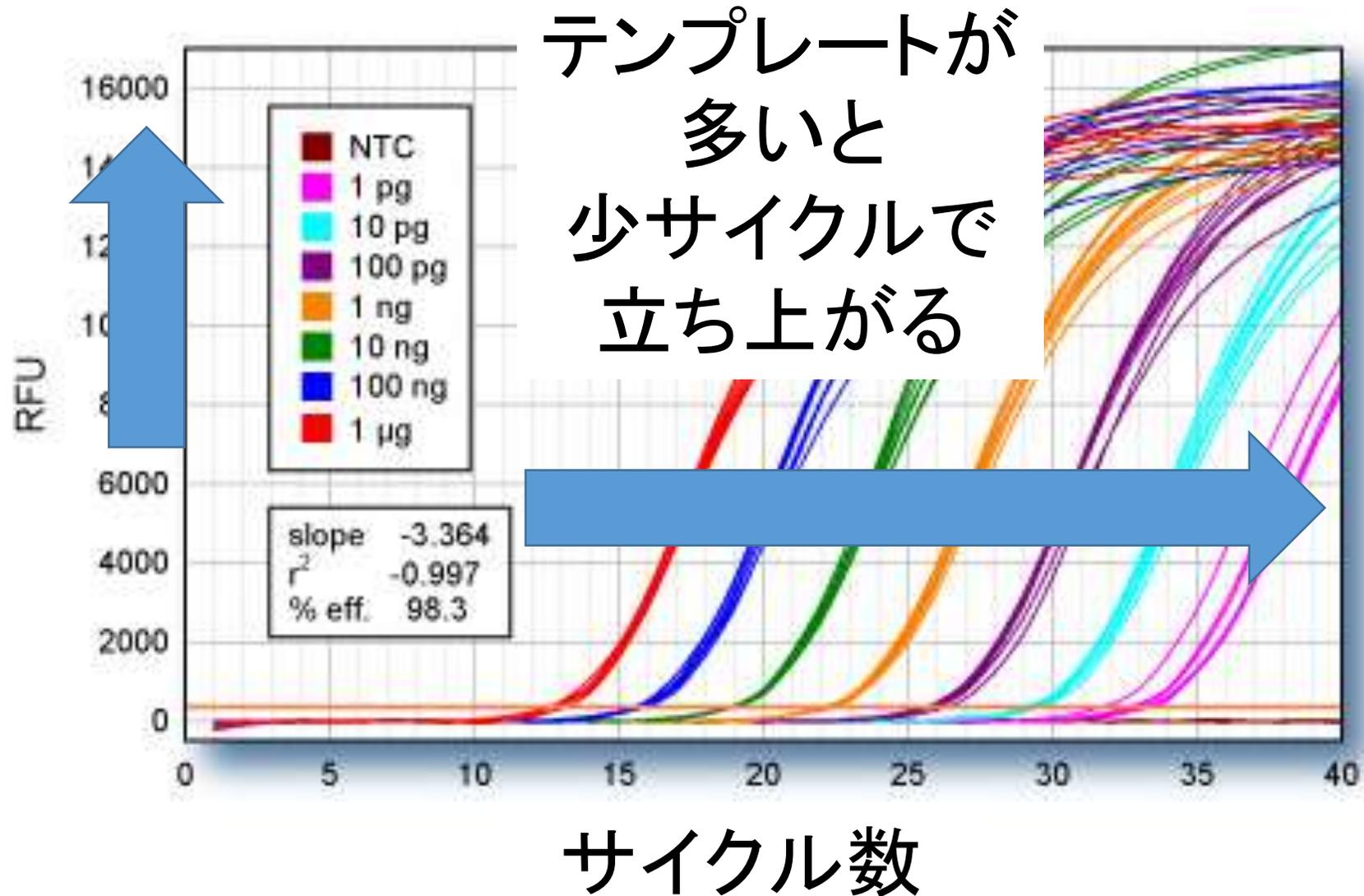
定量的PCR 発現量解析(mRNA)の場合

テンプレート量
X10



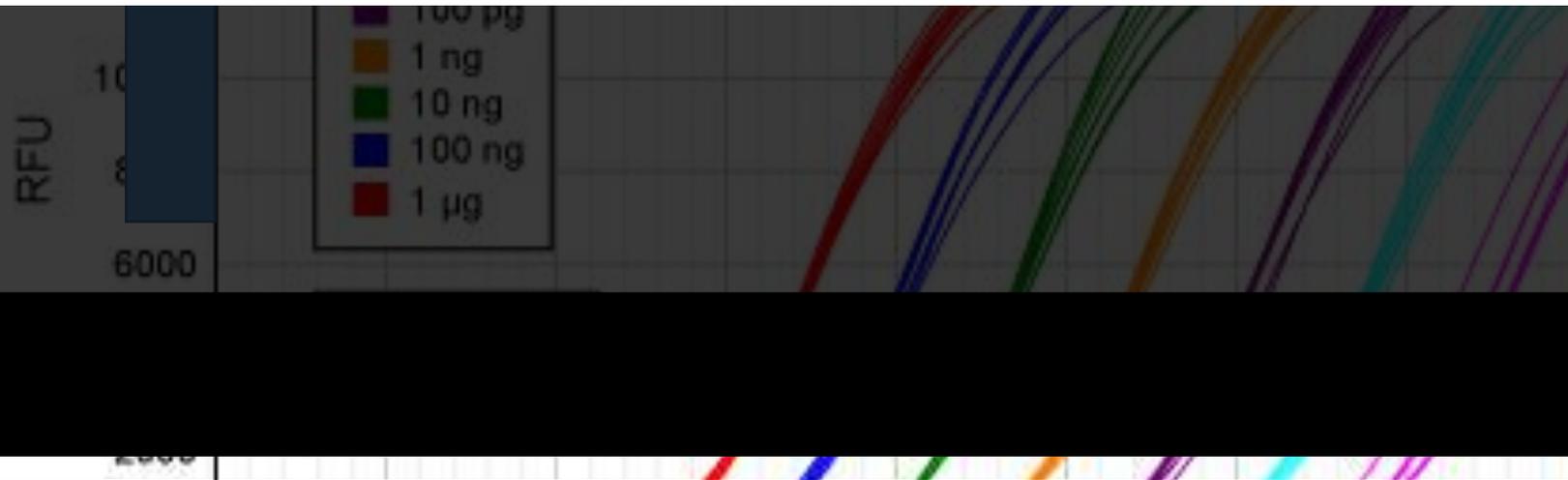
サイクル数

定量的PCR 発現量解析(mRNA)の場合



定量的PCR 発現量解析(mRNA)の場合

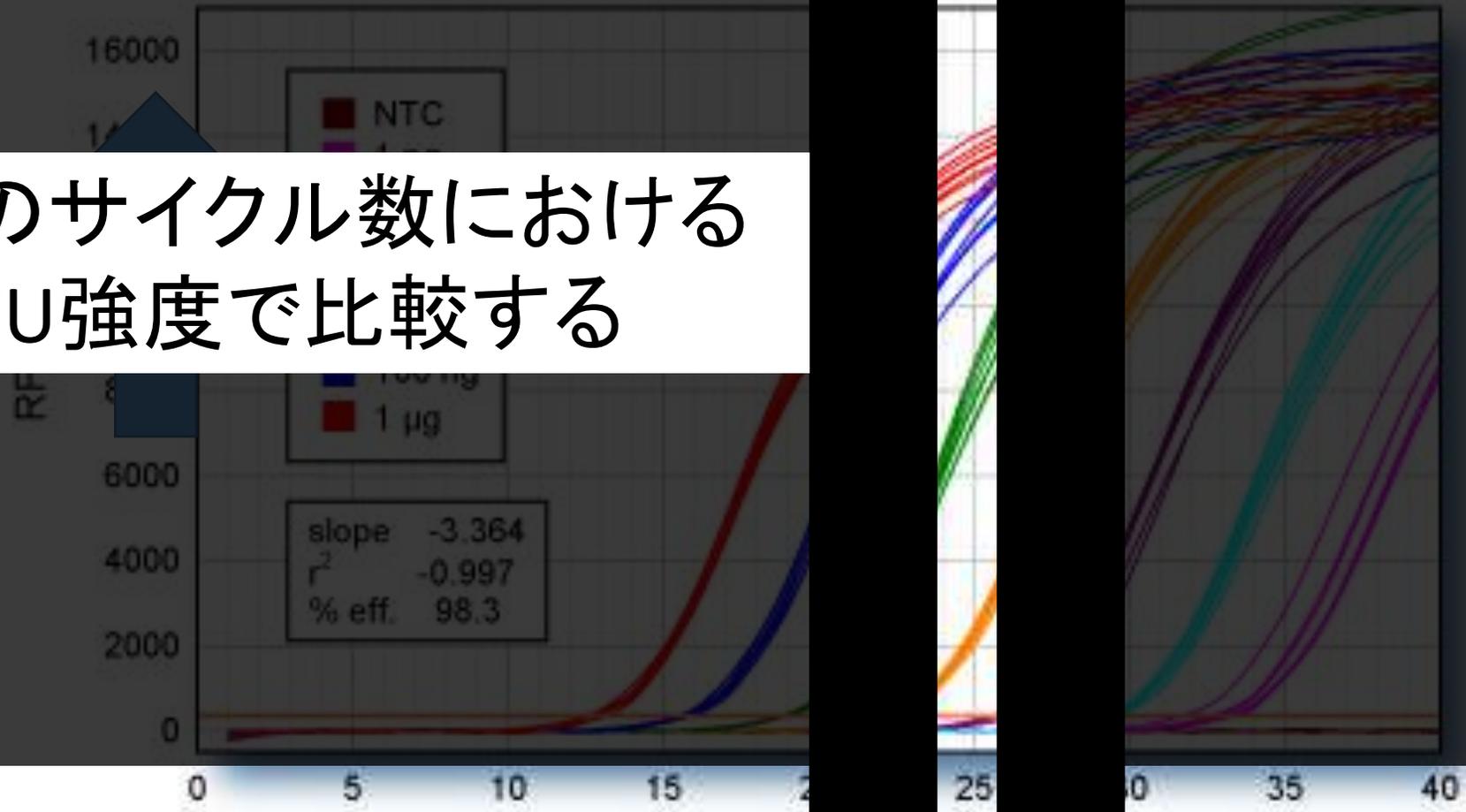
特定のRFU強度に着目して、立ち上がりが必要だったサイクル数を基に、サンプル内のmRNAコピー数を推定する



サイクル数

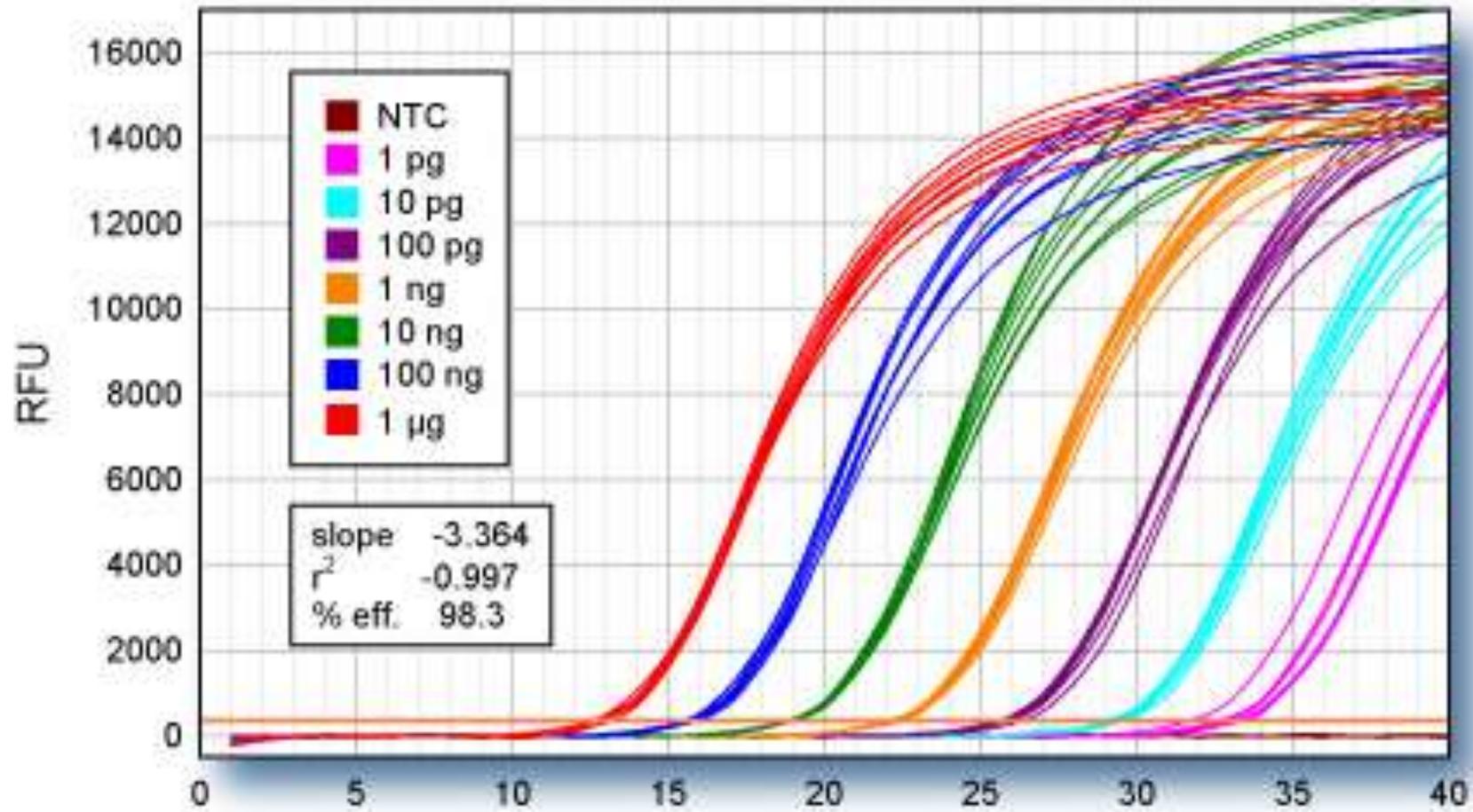
DNA鑑定の場合

特定のサイクル数における
RFU強度で比較する



サイクル数

シグモイドカーブ近似

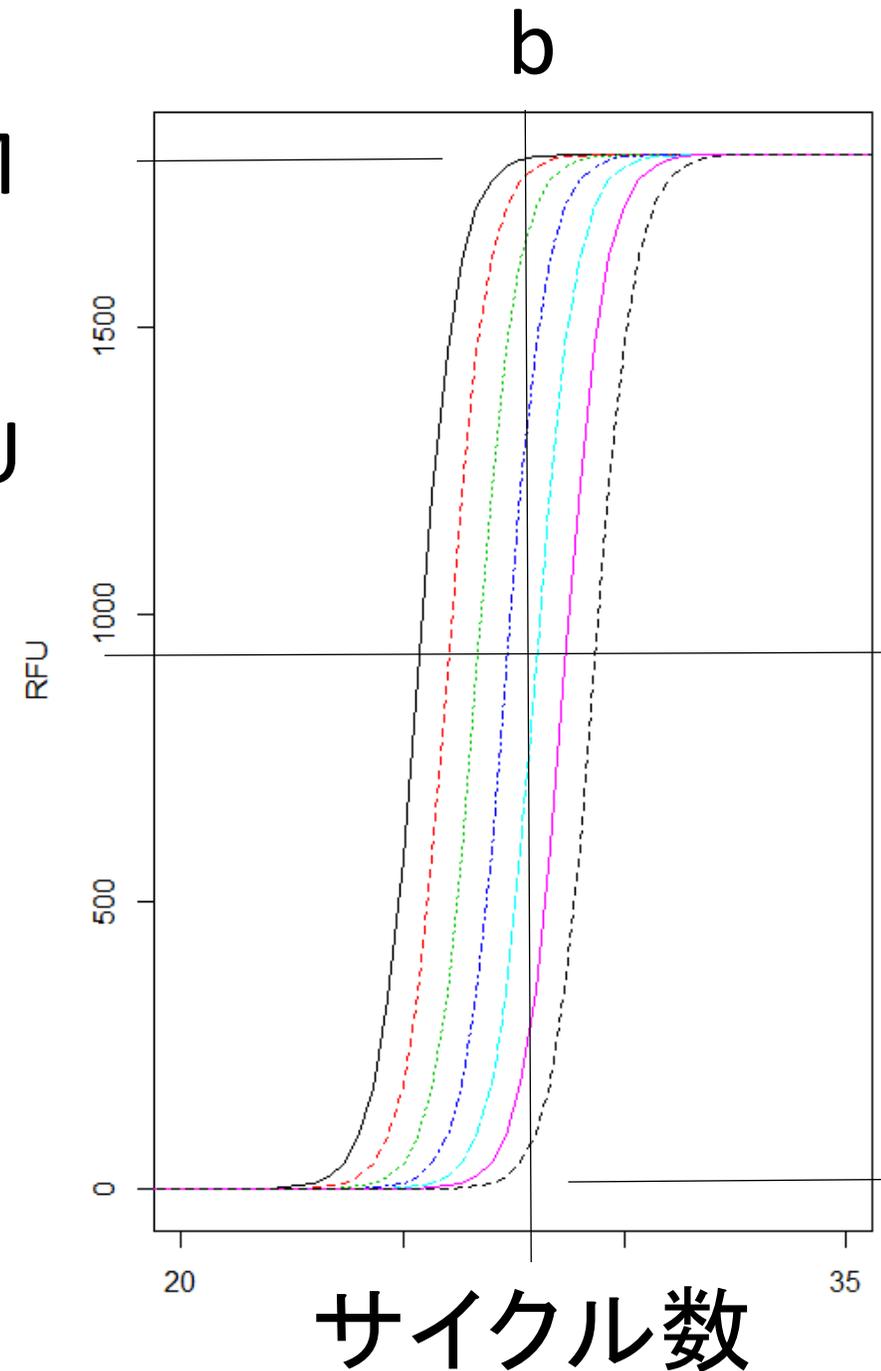


サイクル数

シグモイドカーブ近似

- $RFU = 0 + M/(1+\exp(-k(C-b)))$
- RFUは0を下限、Mを上限とするシグモイドカーブ
- $C(\text{サイクル数})=b$ でRFUの中間値($M/2$)となる
- B の値は、サンプル濃度に依存する。その依存の仕方は指数関数的

M
RFU

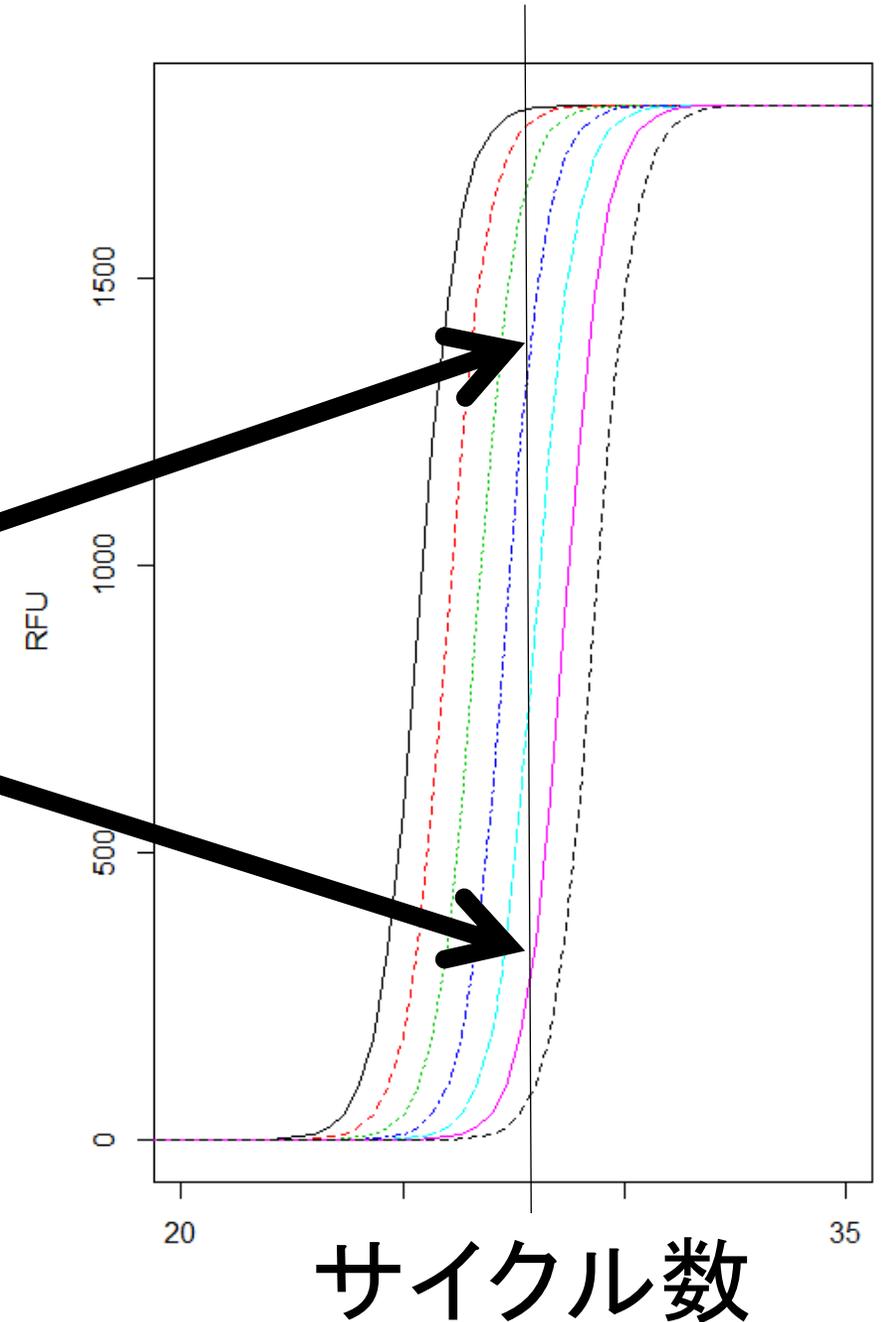


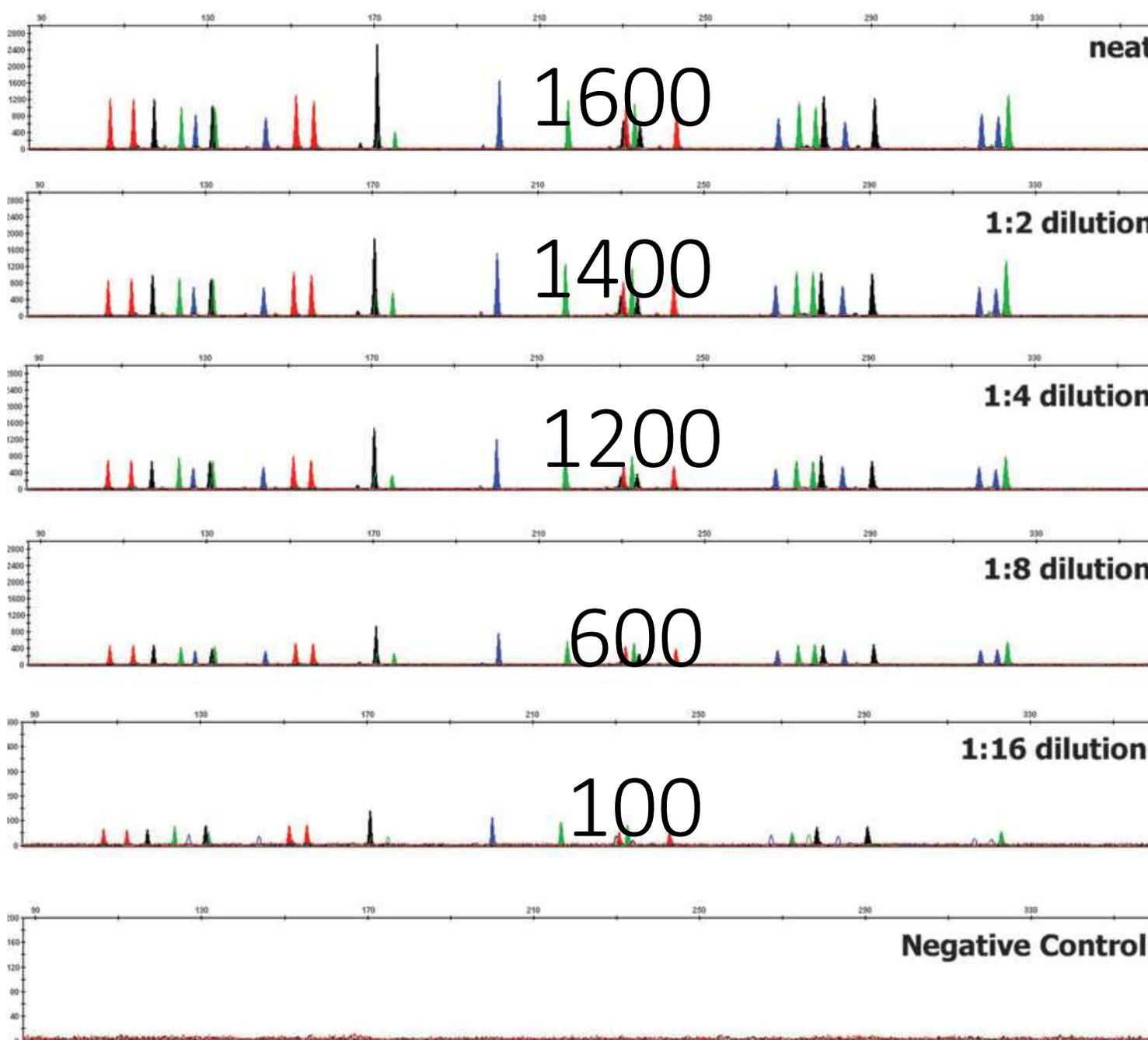
DNA鑑定

特定のサイクル数で切り取る

高濃度サンプルは
高RFU

低濃度サンプルは
低RFU





實際

Development and Validation of the
 AmpF[®]STR
 Identifiler Direct PCR Amplification
 Kit: A
 Multiplex Assay for the Direct
 Amplification of
 Single-Source Samples

<http://onlinelibrary.wiley.com/doi/10.1111/j.1556-4029.2011.01757.x/epdf>

FIG. 5—Effect of varying inputs of WBCs on Indicating FTA[®] cards on peak heights. The results depicted are representative of the 27-cycle amplification on one of the three WBC samples at the indicated input amounts. The data were analyzed with a peak amplitude threshold of 50 RFU.

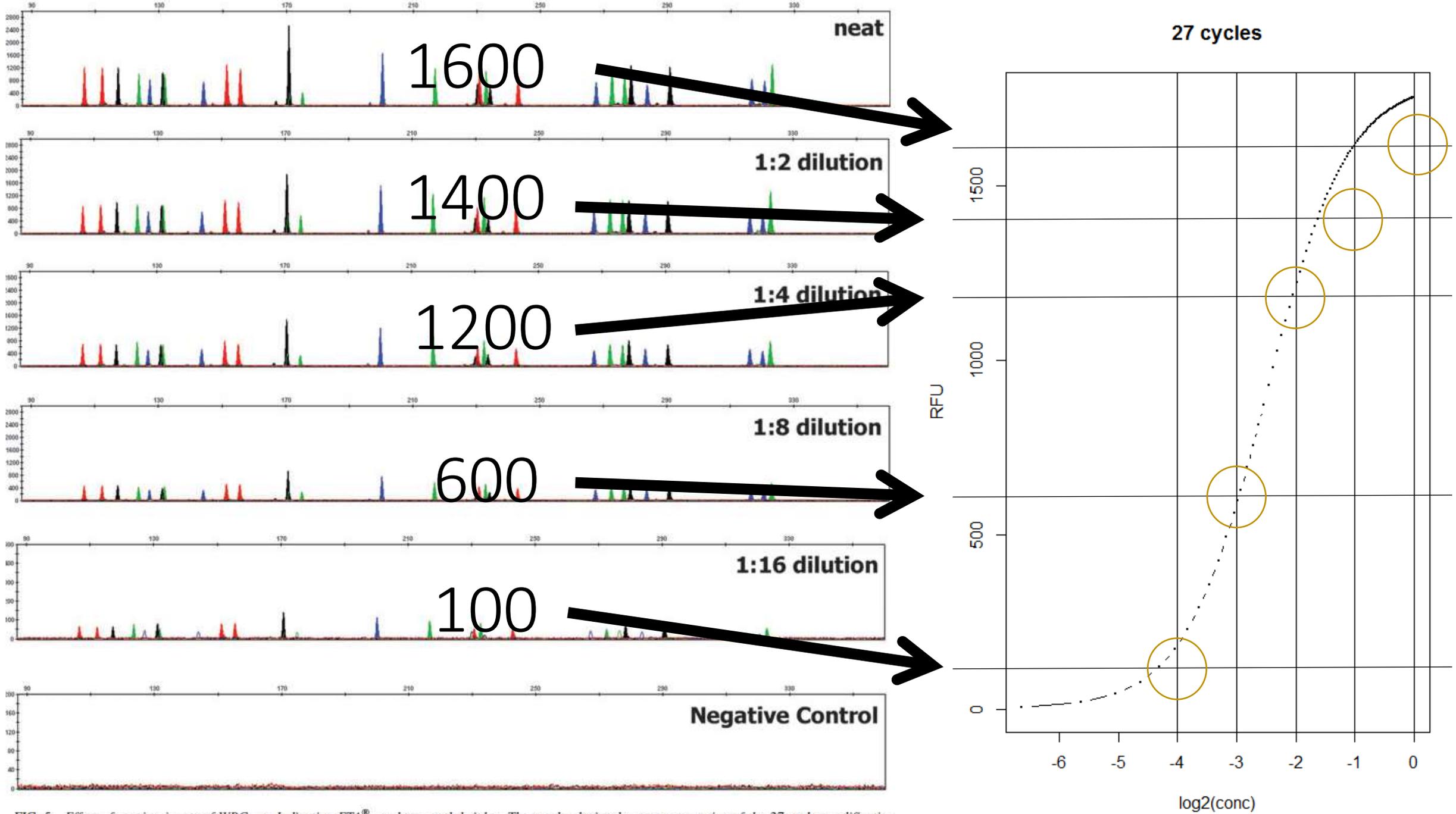


FIG. 5—Effect of varying inputs of WBCs on Indicating FTA[®] cards on peak heights. The results depicted are representative of the 27-cycle amplification on one of the three WBC samples at the indicated input amounts. The data were analyzed with a peak amplitude threshold of 50 RFU.

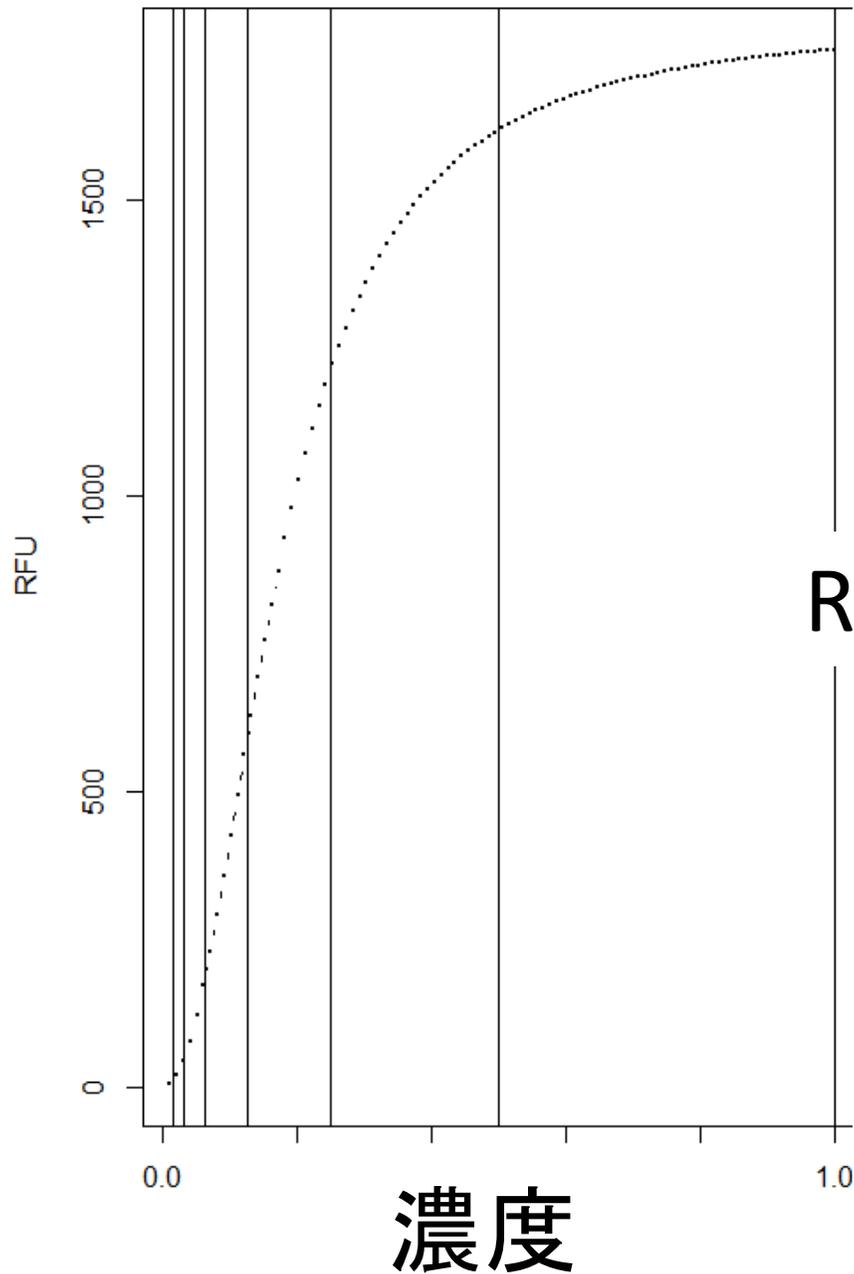
シグモイドカーブ近似も悪くない

もっと近似に凝ってもよいけれど、考え方には違いは出ないので、これでよしとする

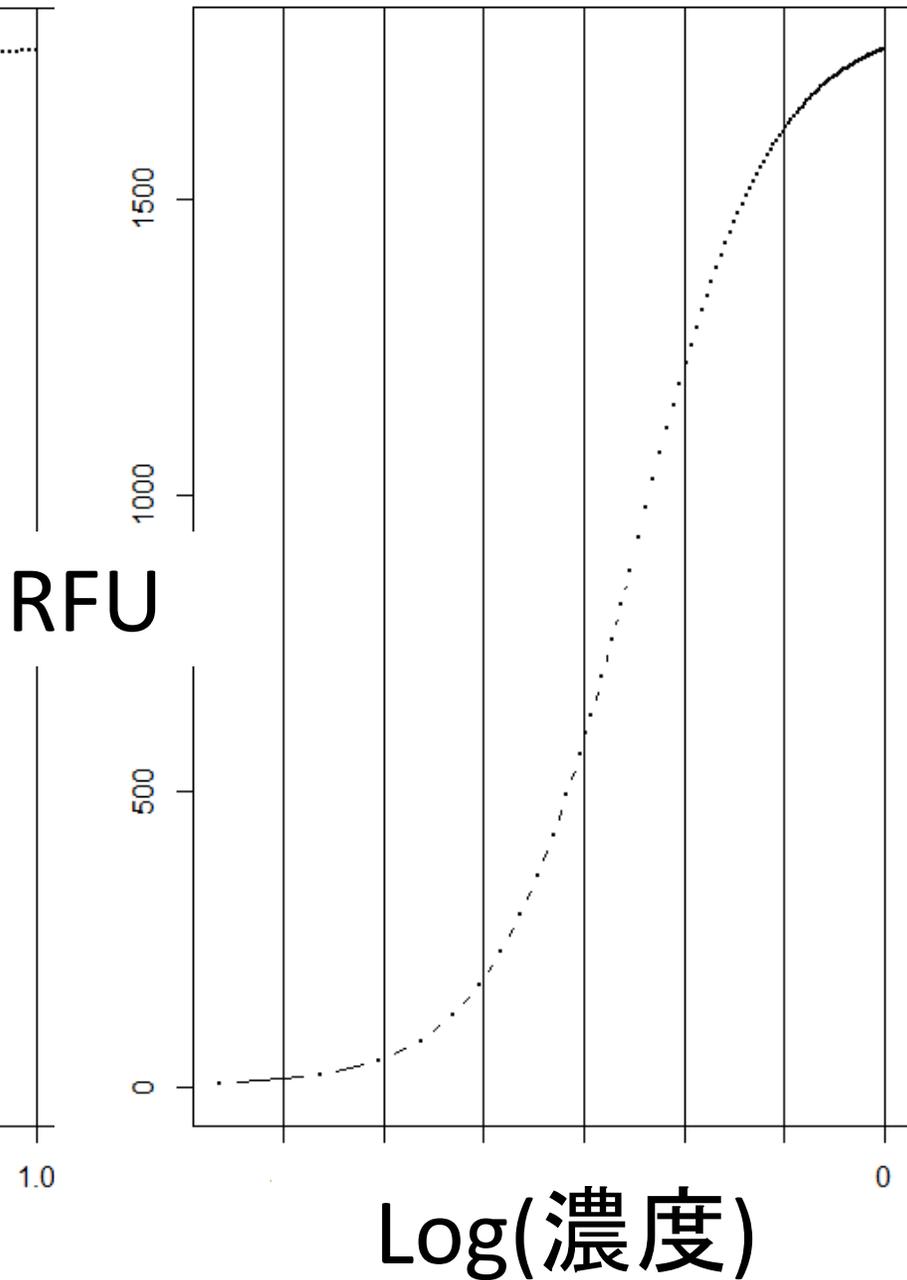
サンプル濃度とRFUの関係

Log(濃度)に対してRFUがシグモイドになる

27 cycles



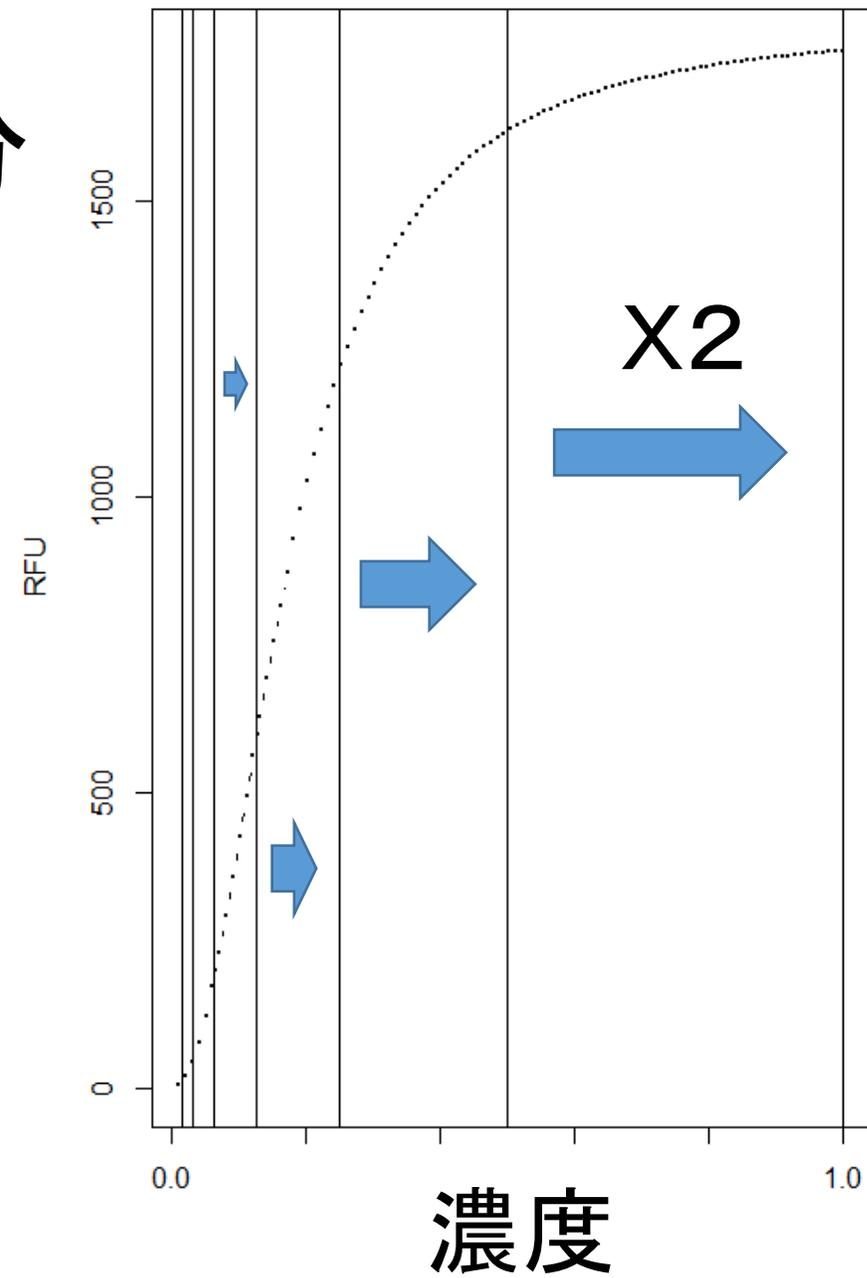
27 cycles



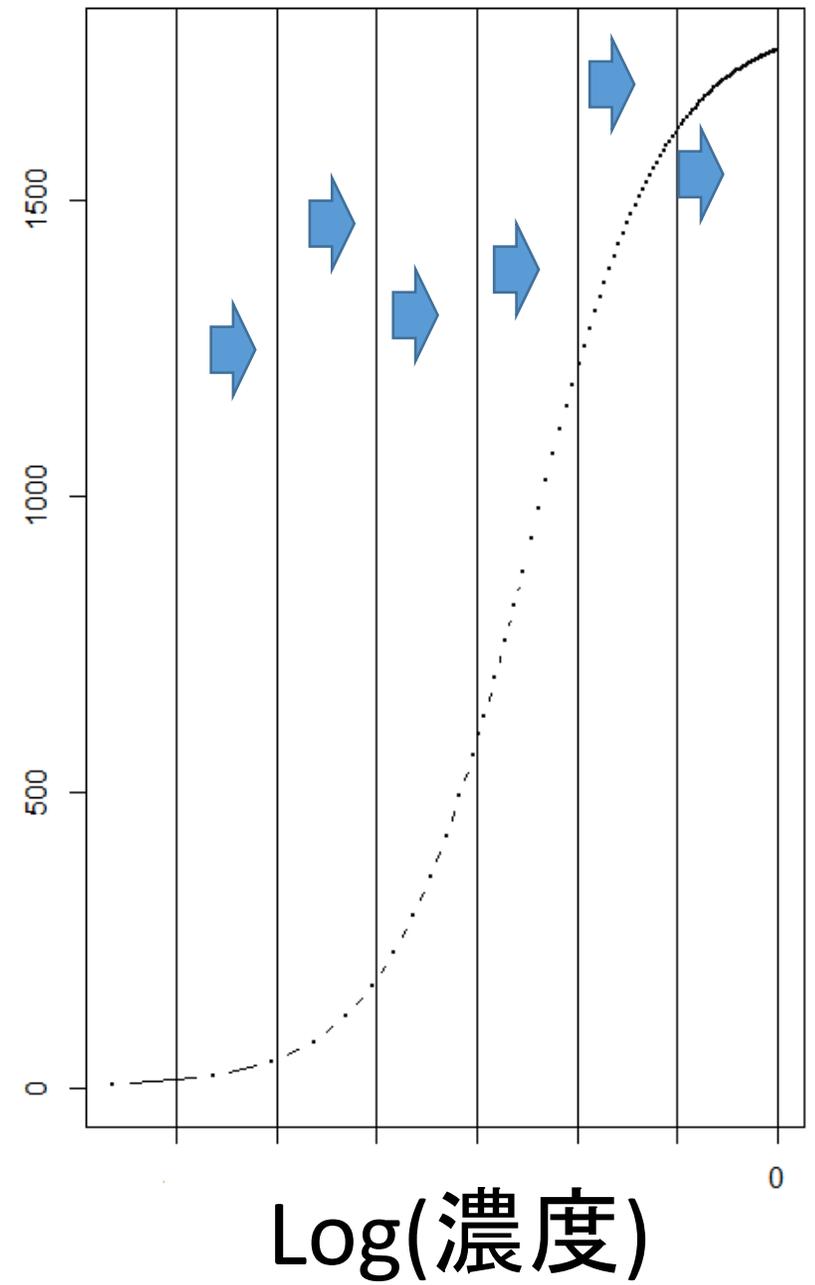
ヘテロ接合体
は濃度が半分

RFUは？

27 cycles



27 cycles

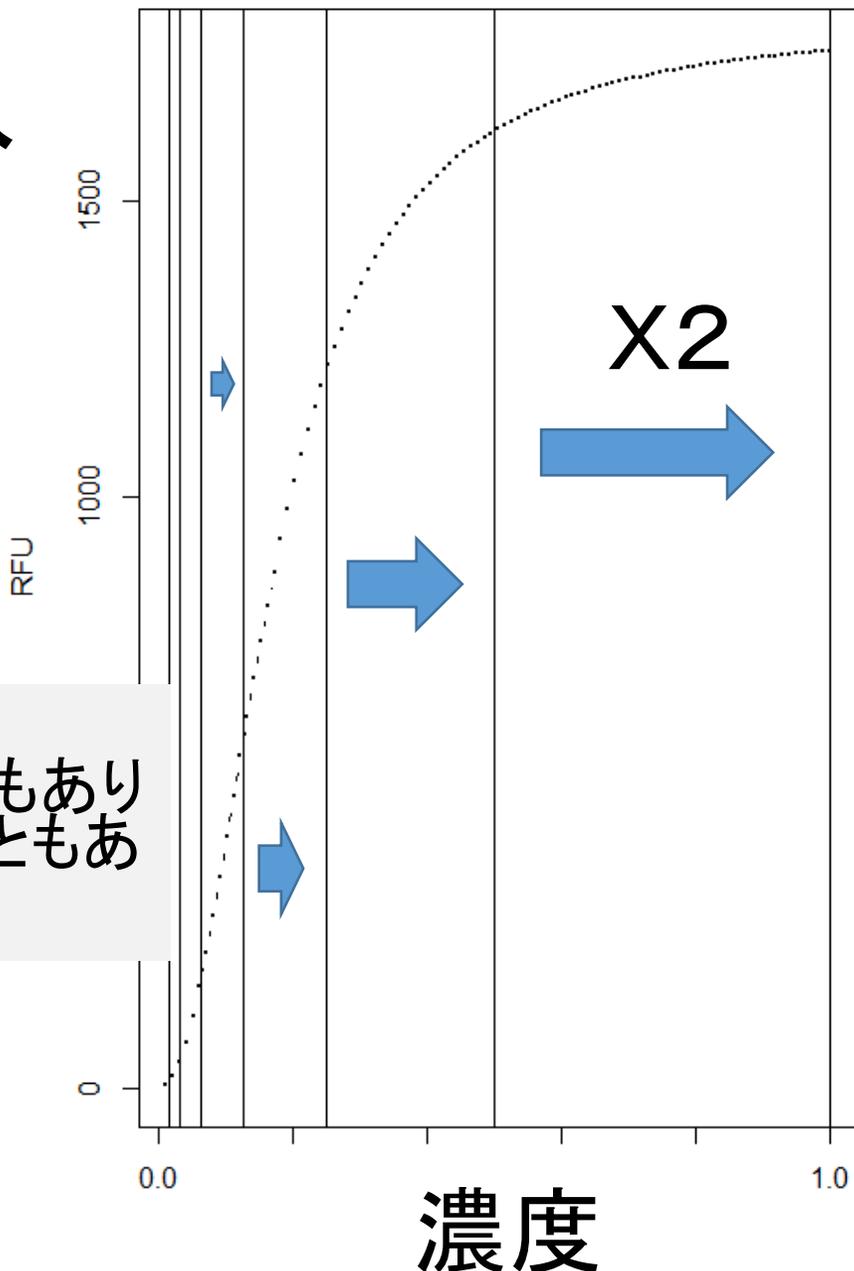


ヘテロ接合体は濃度が半分

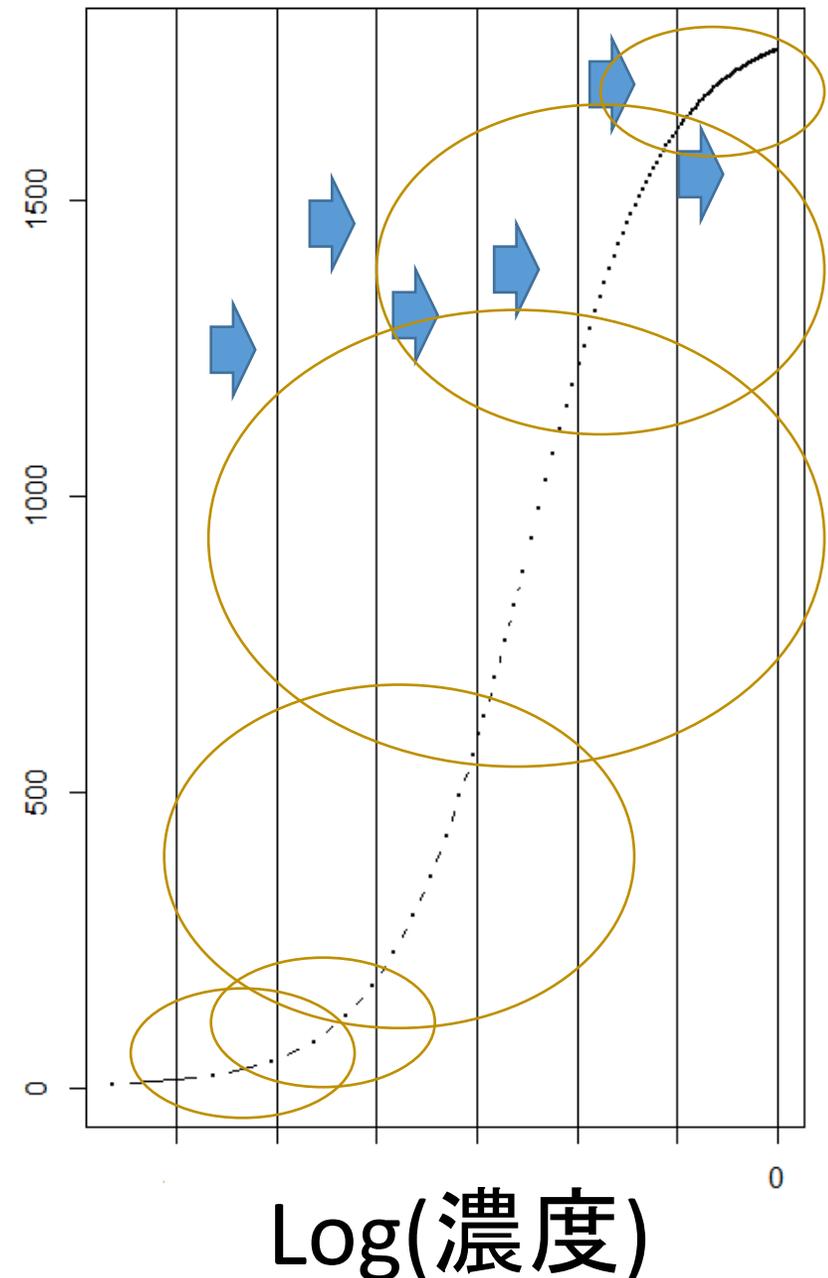
RFUは？

半分になることもあれば
半分より大きめになることもあり
半分よりずっと小さめなこともある

27 cycles



27 cycles



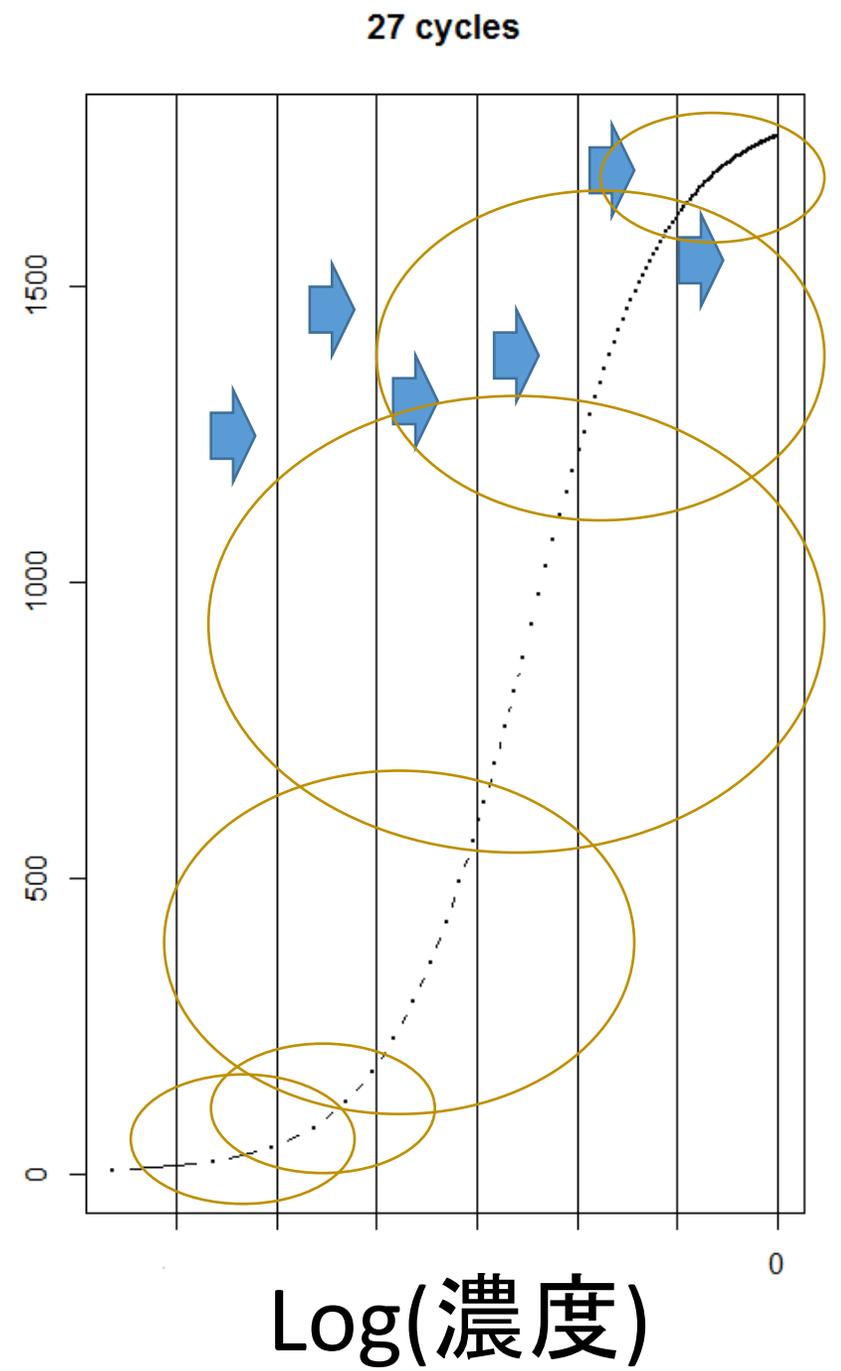
ヘテロ接合体 は濃度が半分

RFUは？

低シグナルでは
ヘテロは低め

高シグナルでは
ヘテロは高め

- 92%
- 75%
- 48%
- 31%
- 25%
- 23%



27 cycles



92%

75%

48%

31%

25%

23%

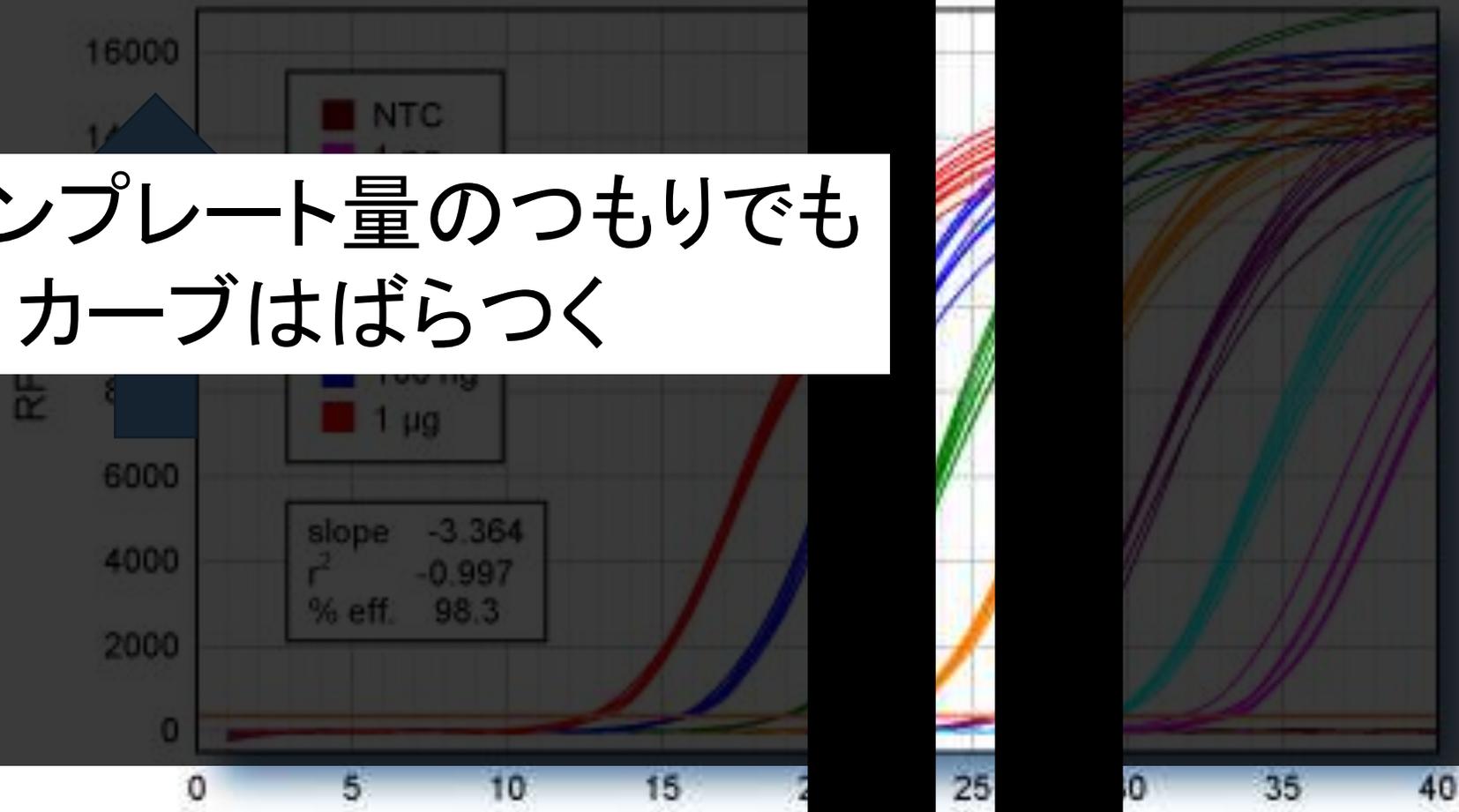
ヘテロ接合体
は濃度が半分

RFUは？

低シグナルでは
ヘテロは低め

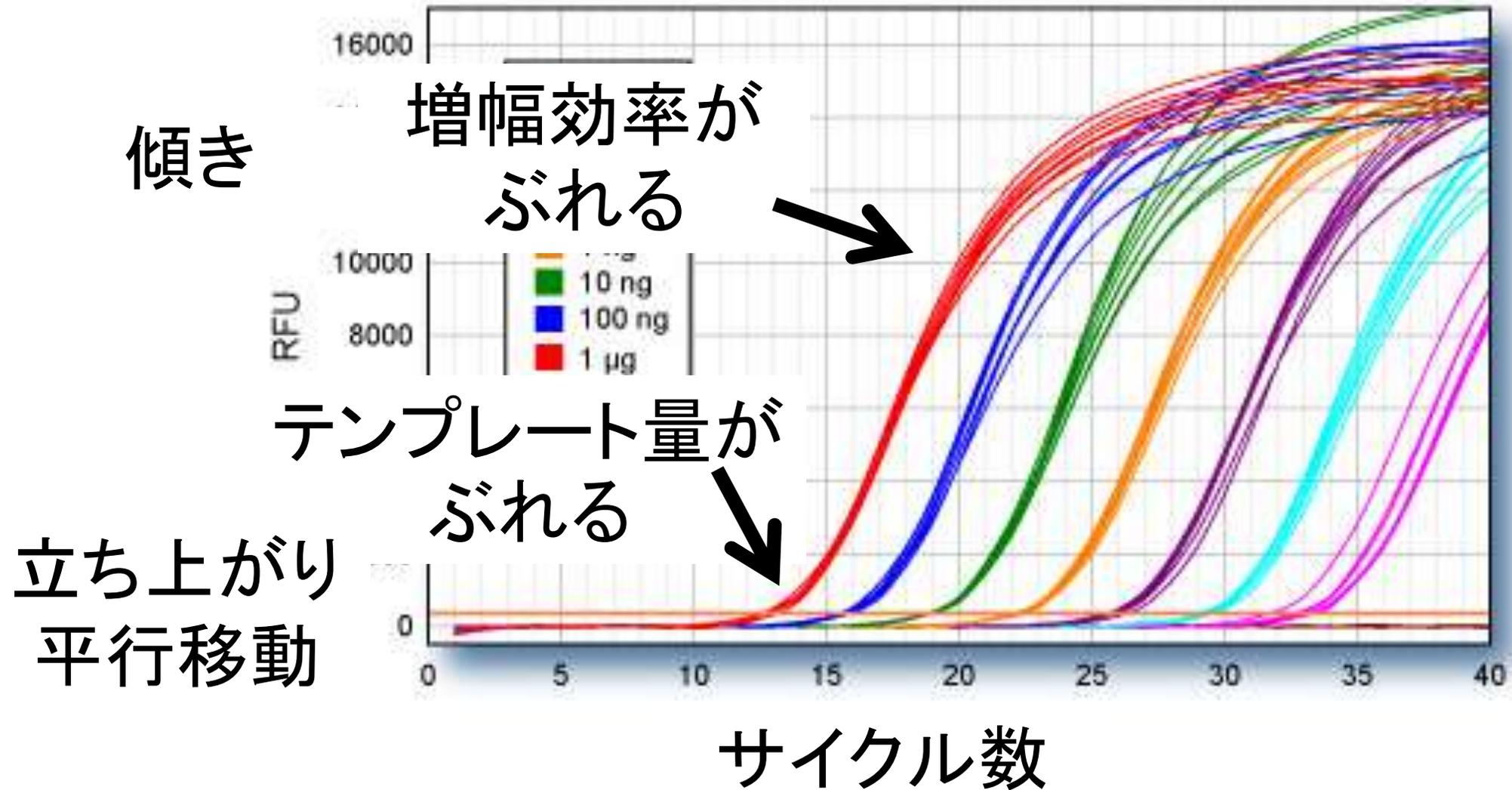
低い方のアレルは
ノイズに埋もれるかも

同じテンプレート量のつもりでも
カーブはばらつく



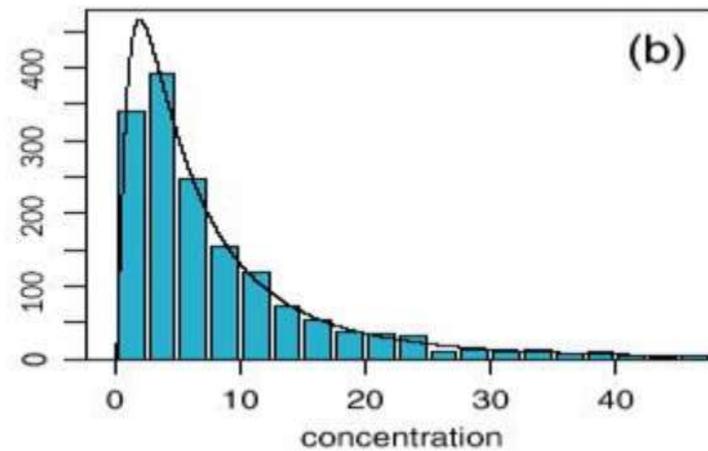
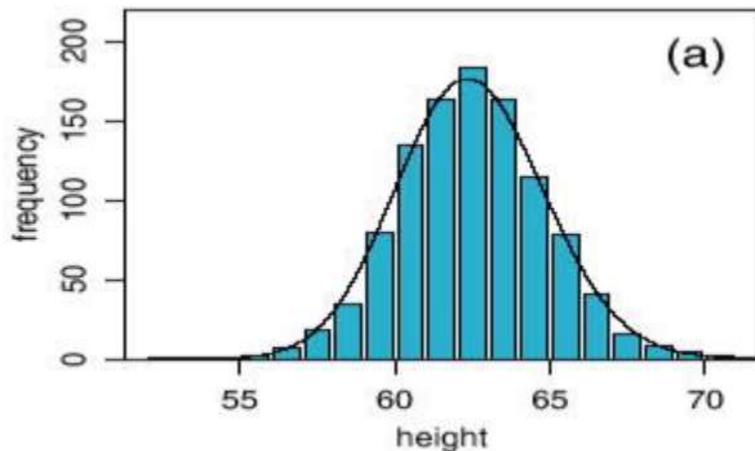
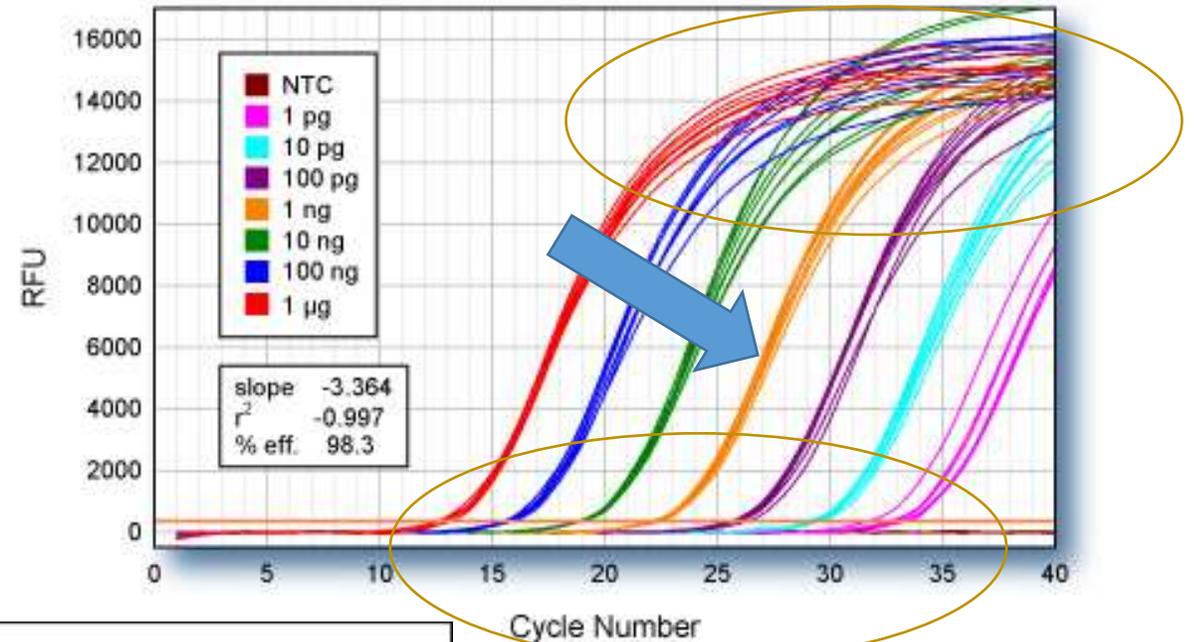
サイクル数

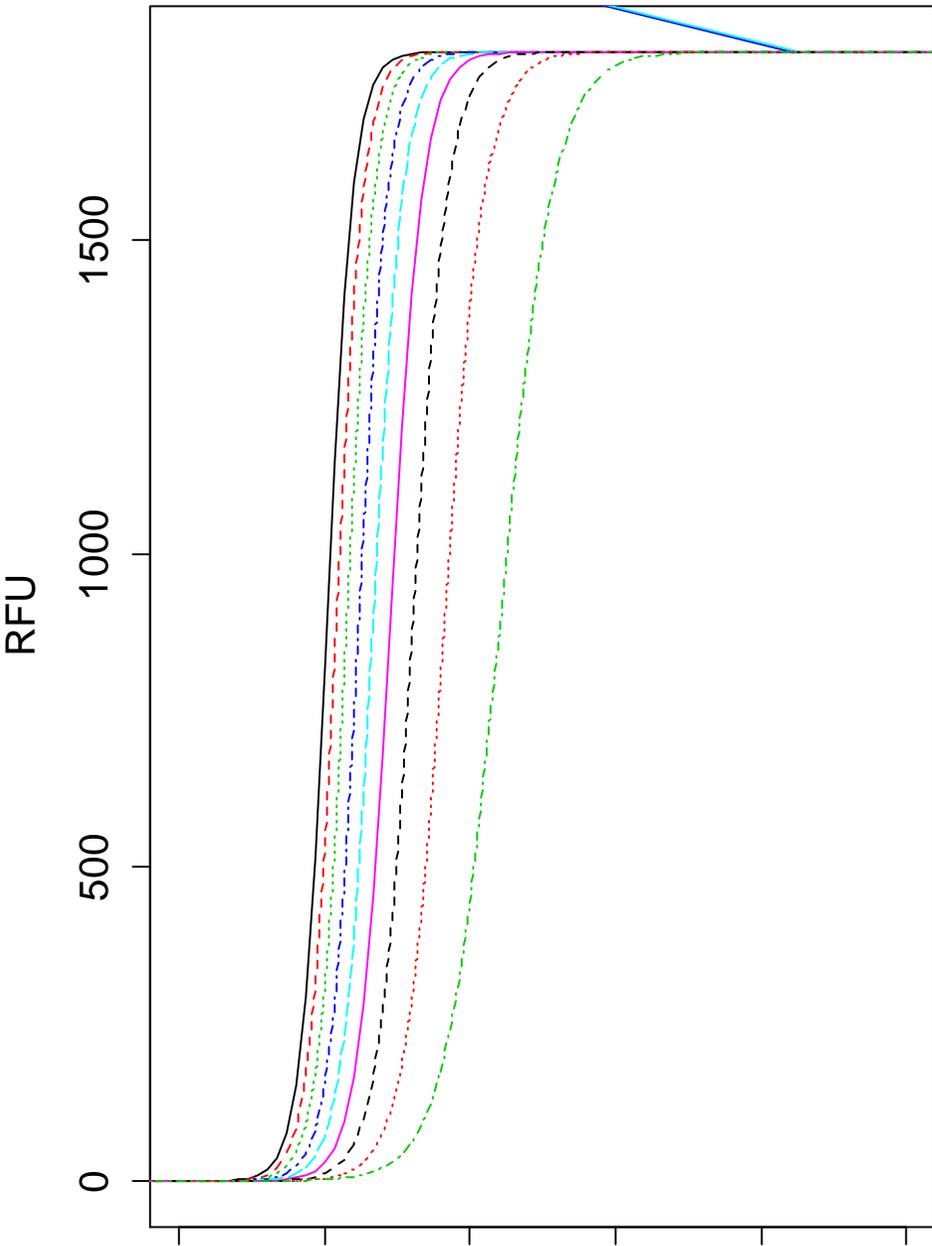
同じテンプレート量のもりでも
カーブはばらつく



バラツキ、エッセンス

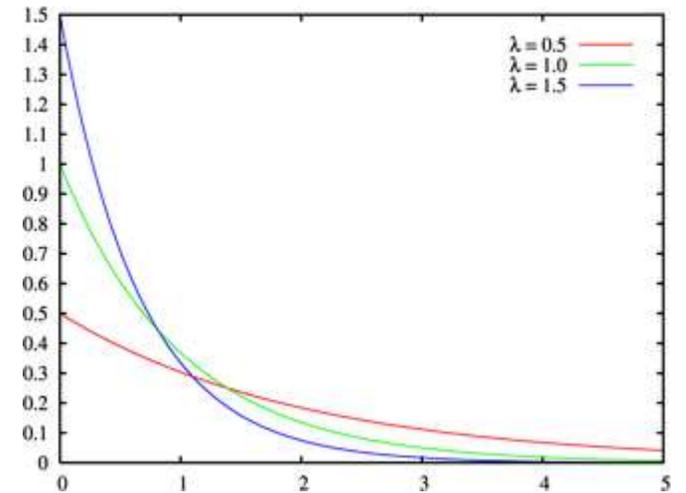
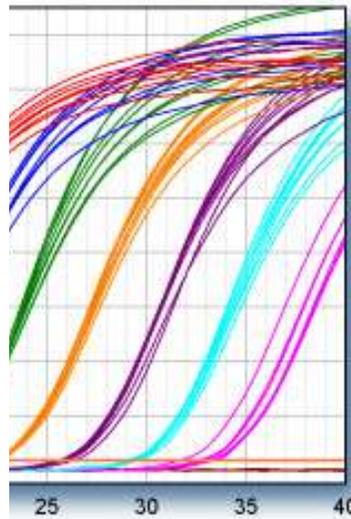
- 何も制約がなければ、大きい方にも小さい方にも均等にばらつく
- バラツキも増幅する
- 制約があると...
 - 窮屈な側には窮屈にばらつく
 - 上に窮屈 下に窮屈
 - ゼロ(正)に窮屈





傾き係数に
指数分布的バラツキがある場合

指数分布

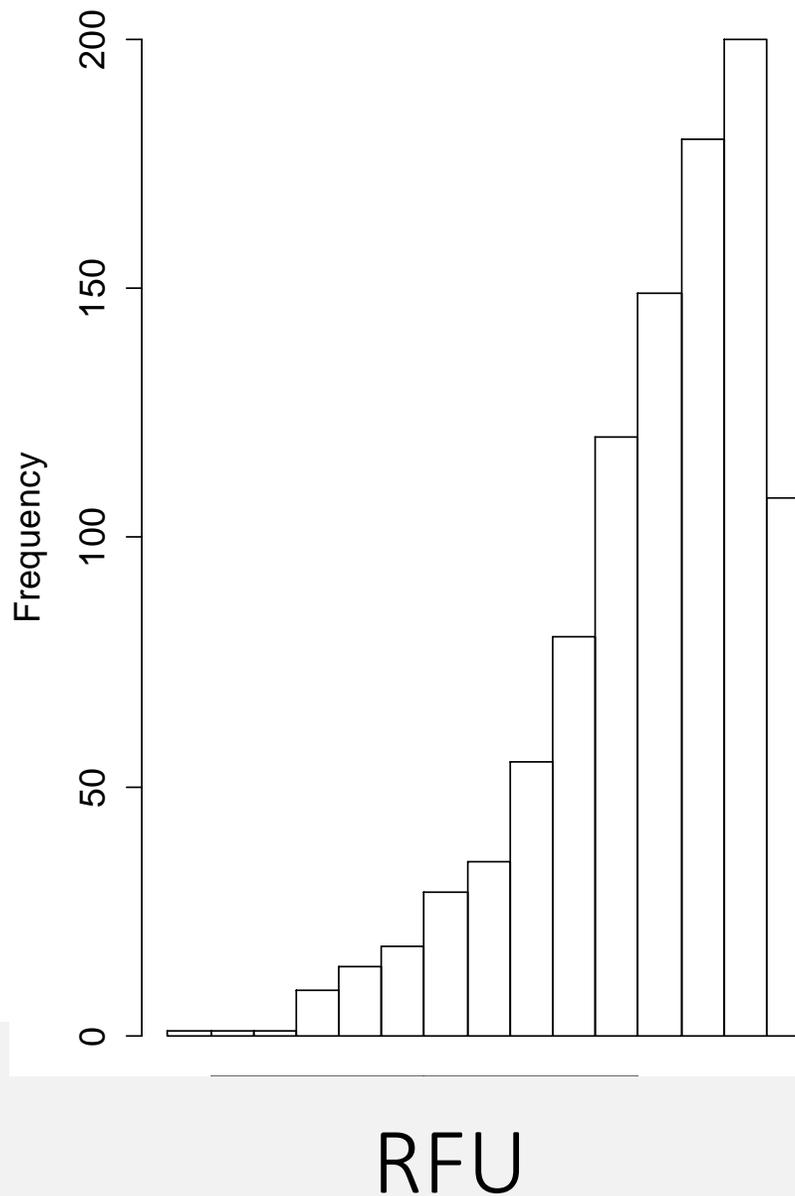


er

傾き係数に
指数分布的バラツキがある場合

RFU分布は
下広がり

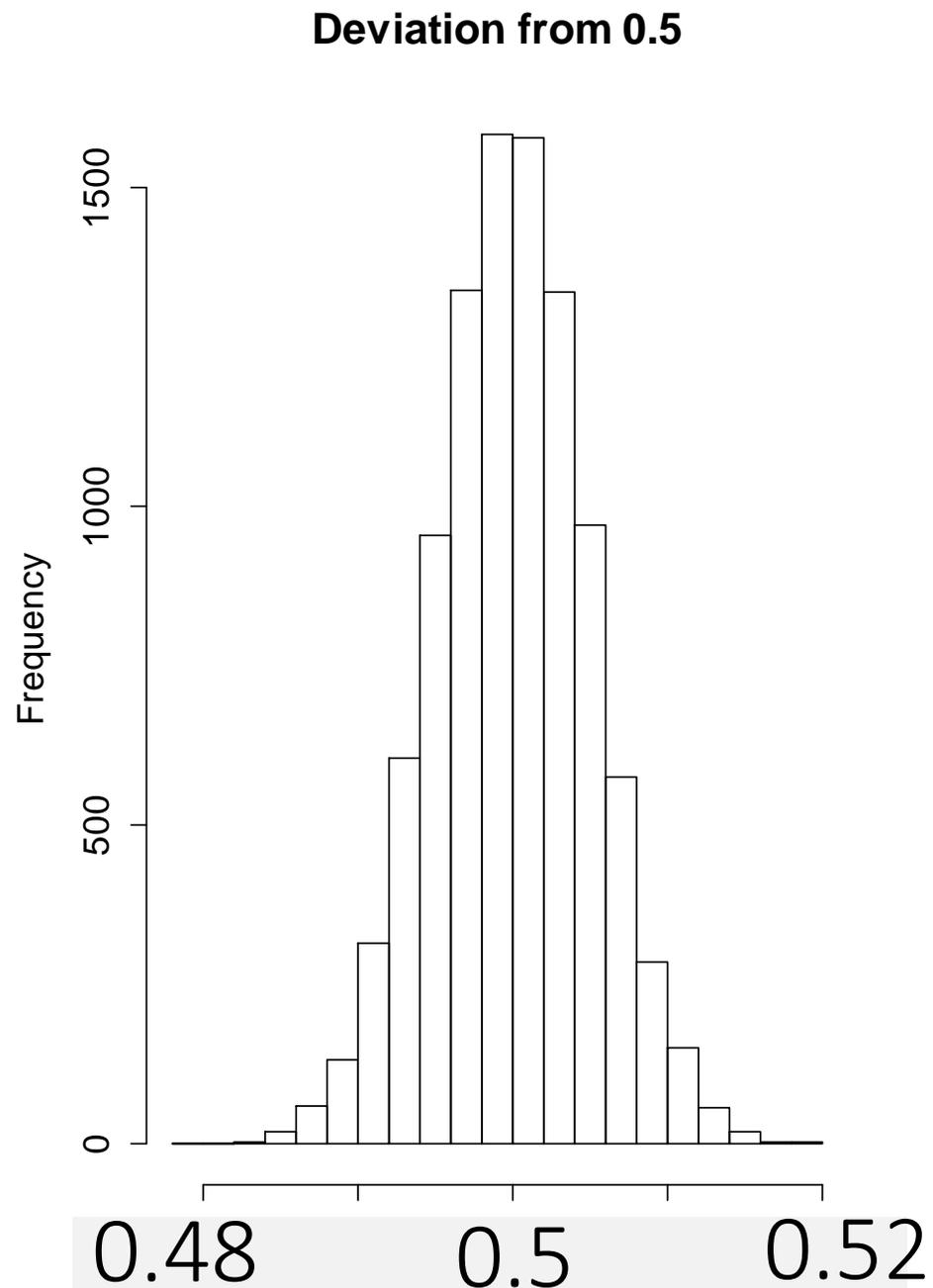
頻度



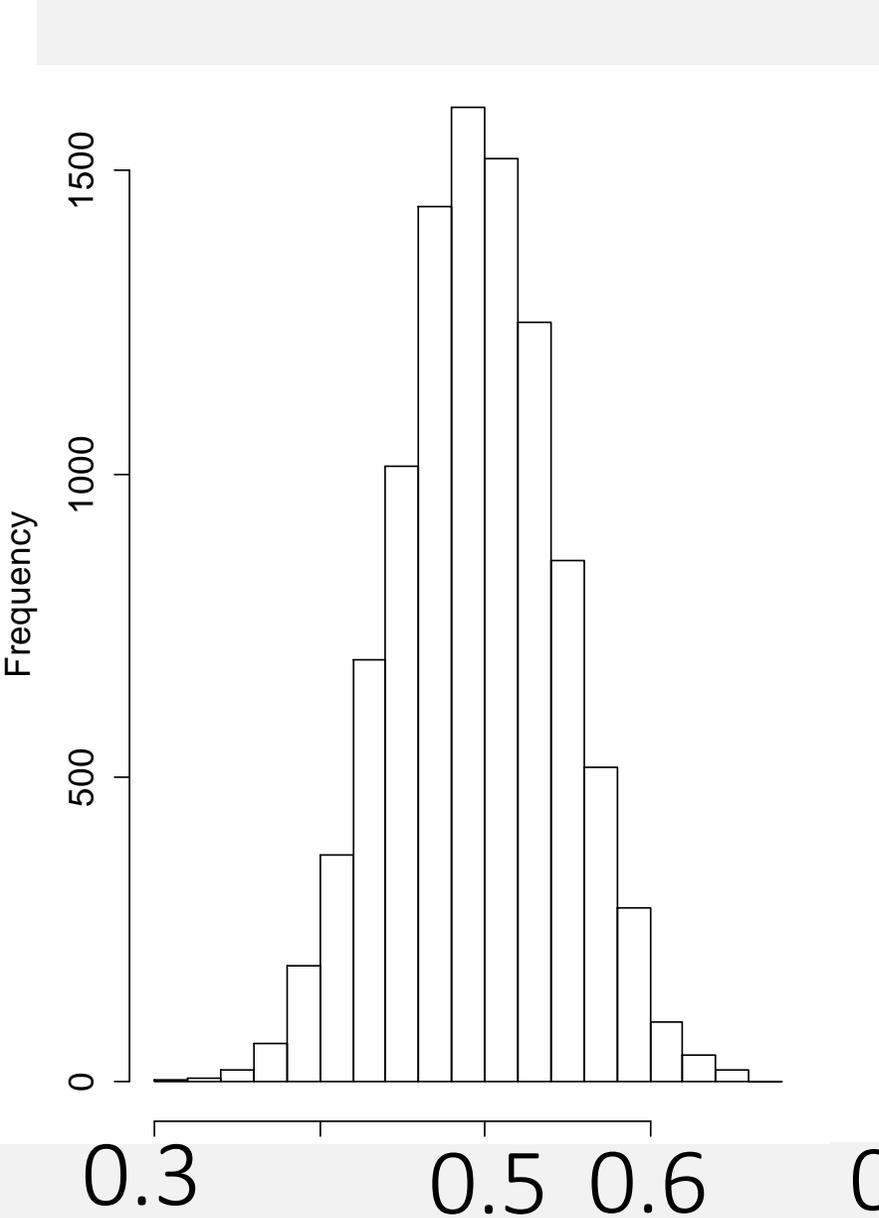
最後に二項分布

ヘテロ接合体の2アレルのサンプル中のコピー数もばらつく

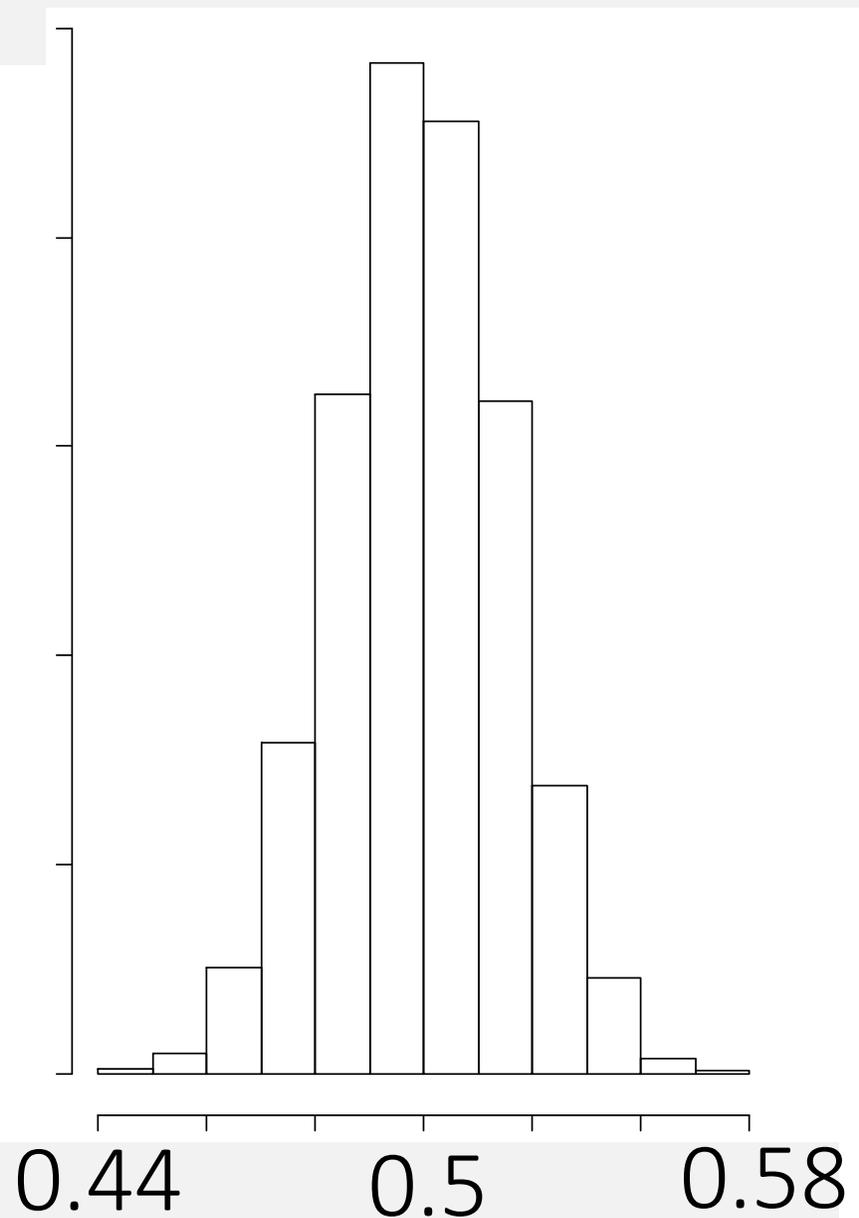
- 25-100 ng のヒトゲノムDNAに10,000強の遺伝子コピーがある
- 1ml の血液から細胞・ゲノムDNAの回収率が10%として、試料を作り、そこから1/100を実験用に調整したとすると...



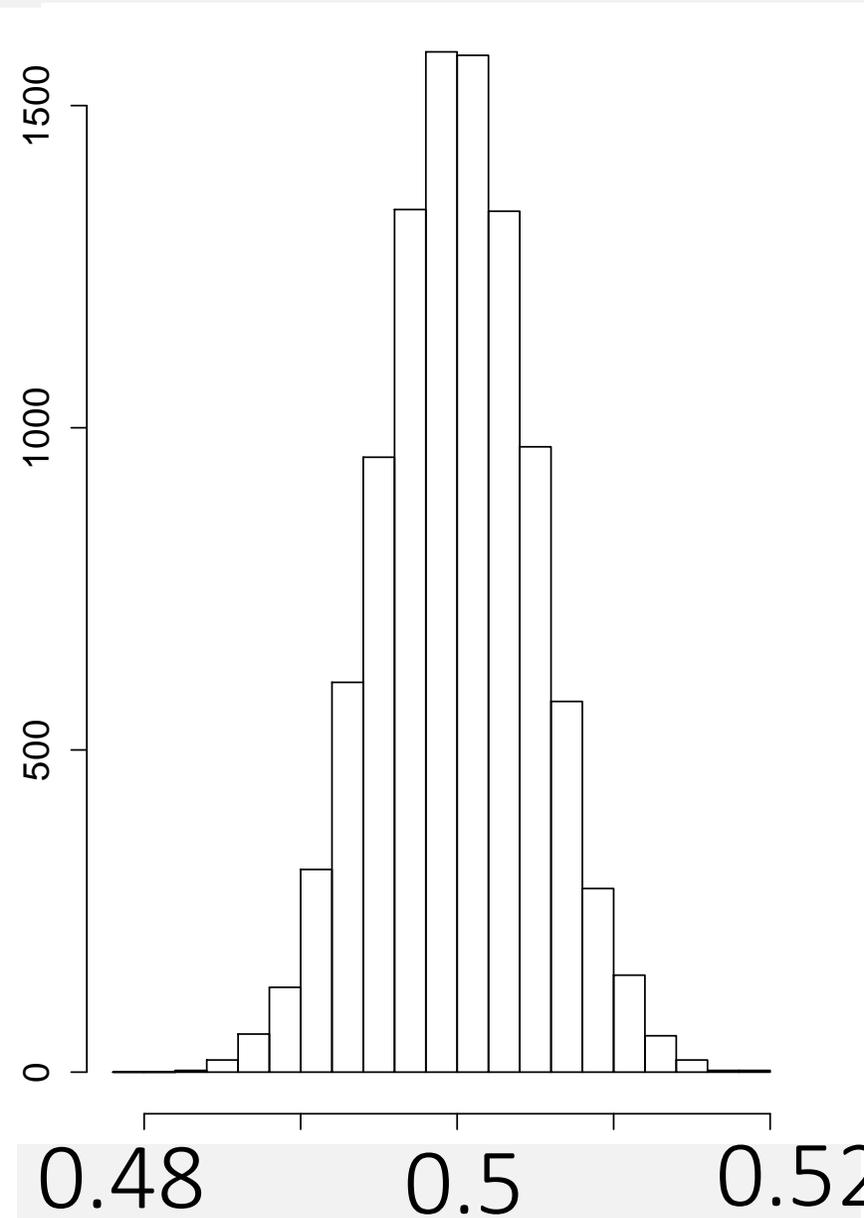
コピー数 100



1000



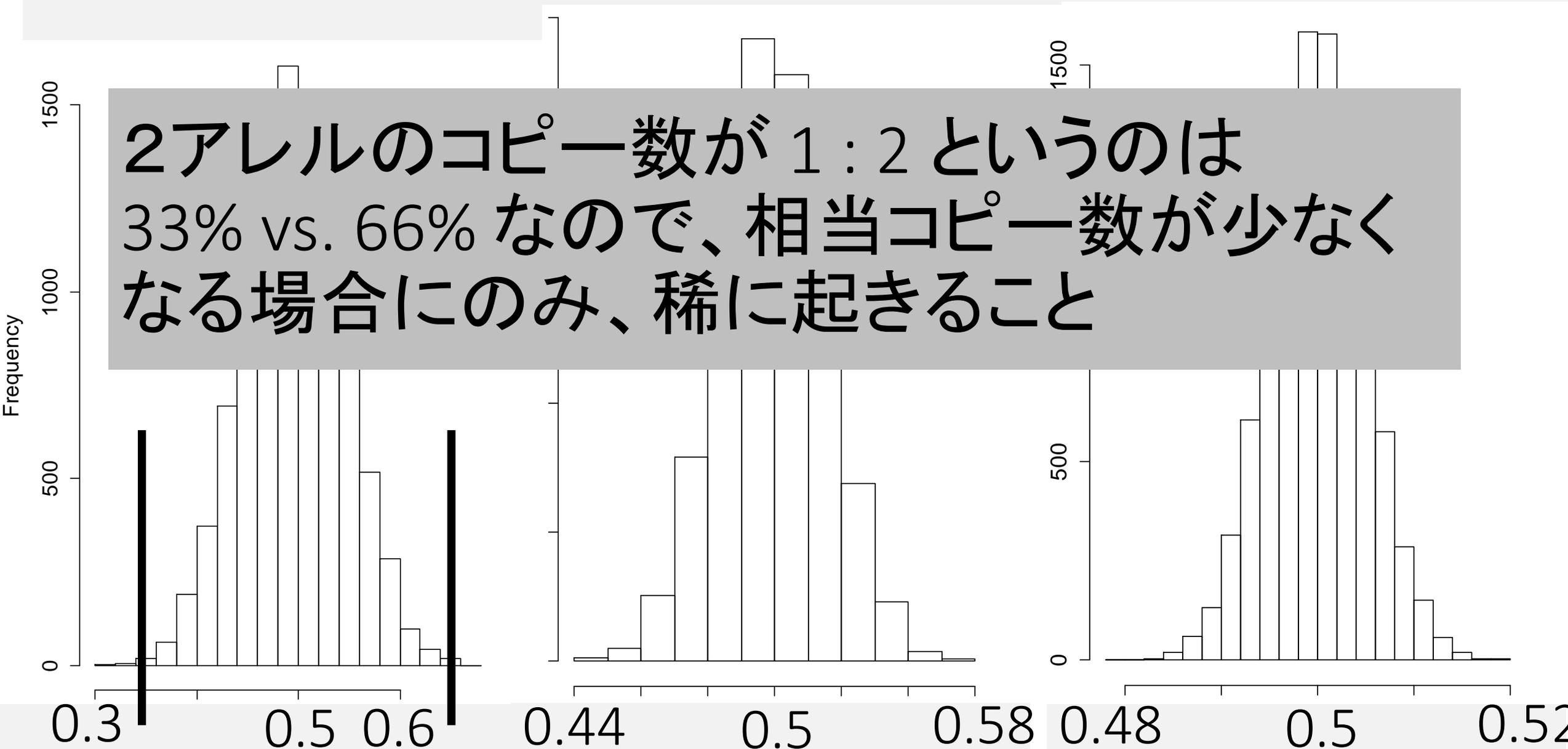
10000



コピー数 100

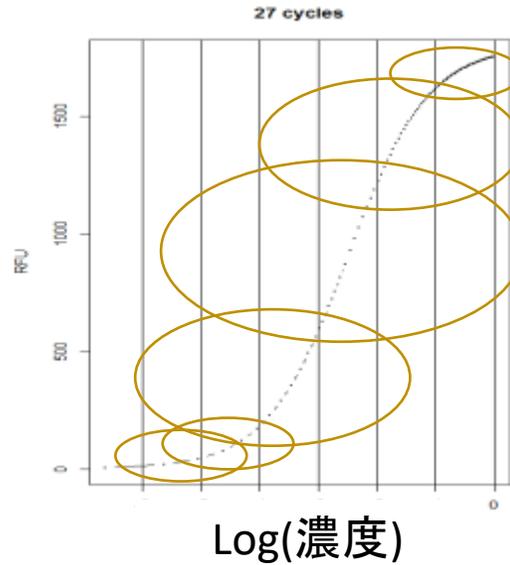
1000

10000

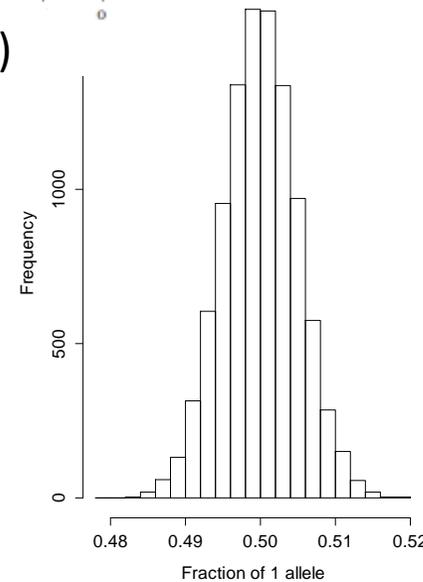


まとめ

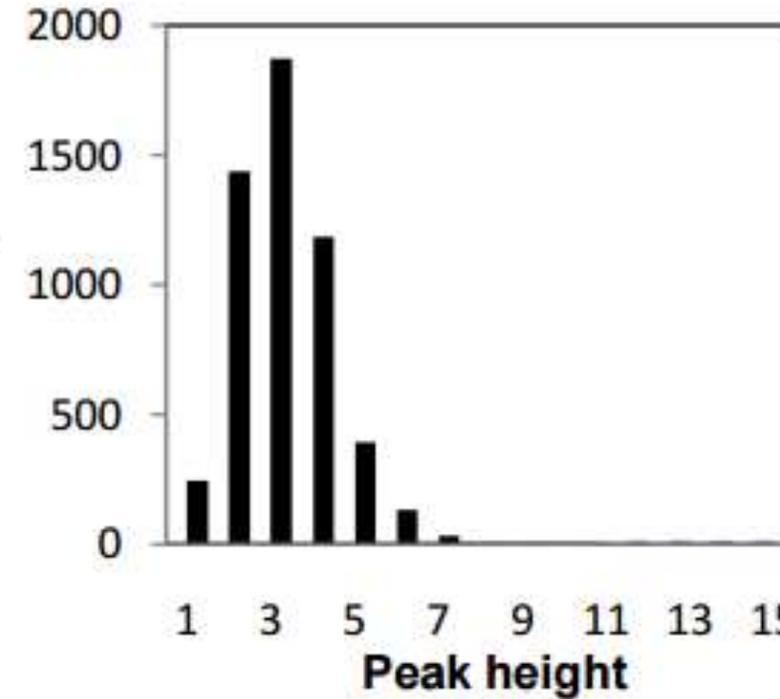
- RFUシグナル
 - ノイズとシグナル
- ノイズ
 - ポアソン対数正規分布
- 真のシグナル
 - ヘテロ2アレルのバラツキ: 二項分布
 - ~2倍
 - 増幅のバラツキ: 指数分布
 - 低RFUに裾を引く
 - コピー数~RFU関係: シグモイド曲線
 - 低シグナルでは、1/2コピーのRFU高は1/2より小さい



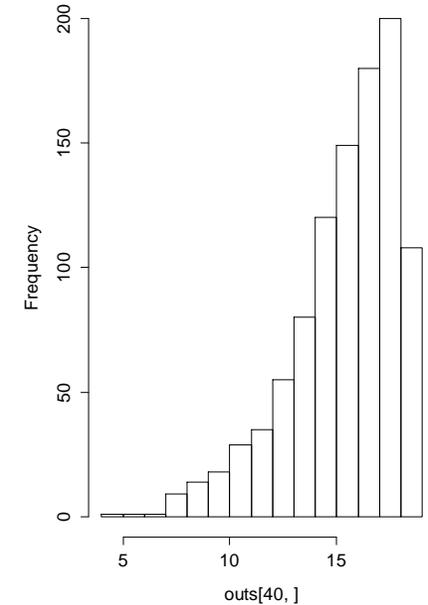
Deviation from 0.5



ber of peaks

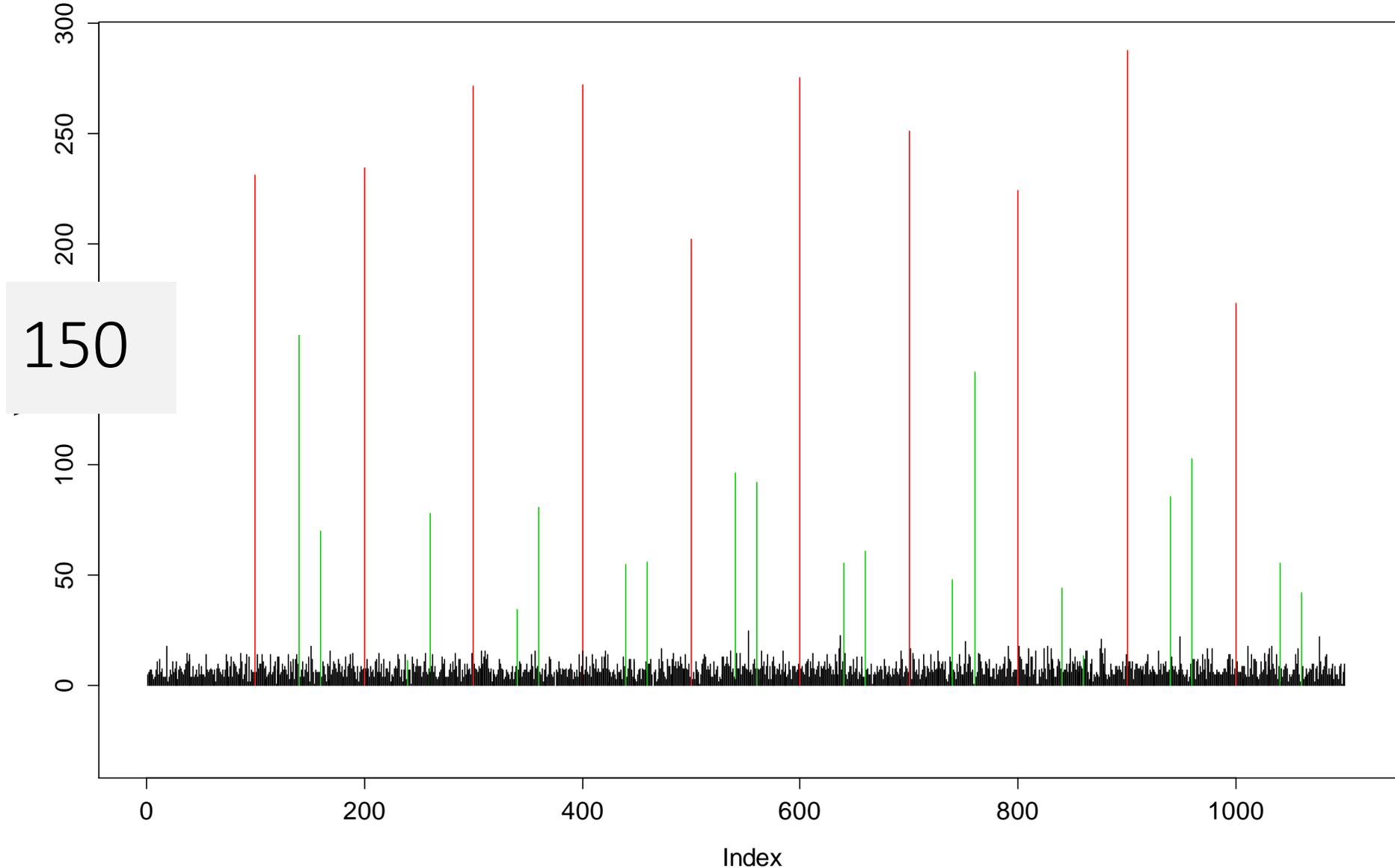


Histogram of outs[40,]



ノイズ・真のシグナルシミュレーション

- ホモの真のシグナルは150RFUはある
- ヘテロは二項分布でばらつく
- 増幅傾斜のバラツキは指数分布
- ノイズはポアソン対数正規分布



その他の参考資料・文献

- In general
 - http://www.nfstc.org/?dl_id=26
- Peak Height Ratio (Heterozygous)
 - http://www.nist.gov/forensics/upload/04_Final-2372A-Dilution-Series-080314J.pdf
 - <http://www.ncbi.nlm.nih.gov/pubmed/24799164>
- Low Template
 - <http://www.promega.jp/resources/profiles-in-dna/2010/scientific-issues-with-analysis-of-low-amounts-of-dna/>
 - <http://www.cstl.nist.gov/strbase/LTDNA.htm>

スライド・シミュレーションRコード文書へのアクセスは分野FacebookにURLを掲載しています



statgenetKyoto

